DOCUMENT RESUME

ED 192 169 CE 027 164

AUTHOR Block, Arthur R.; Rebell, Michael A.
TITLE The Assessment of Occupational Competence. 5. Competence Assessment and the Courts: An Overview of the State of the Law.
INSTITUTION McBer and Co., Boston Mass.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Feb 80
CONTRACT 400-78-0028
NOTE 138p.; For related documents see CE 027 159-165.

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS Access to Education: Adults; *Certification; *Competence; Compliance (Legal); *Court Litigation; *Educational Testing; Employment Practices; Employment Qualifications; Higher Education; Job Skills; Minimum Competency Testing; *Occupational Tests; Postsecondary Education; *Selective Admission; Standards; Teacher Certification
IDENTIFIERS Civil Rights Act 1964; *Occupational Competence Assessment

ABSTRACT

One of seven sections of a report that examines the assessment of occupational competence, this chapter provides a review of how the courts have responded to assessment practices. The difference in the perspectives held by the users and originators of assessment procedures and those held by the courts is developed by examining key court cases which followed upon the enactment of Title VII of the 1964 Civil Rights Act. In light of the courts' present posture, the implications of this difference in perspective are discussed for selection, licensing and certification, and educational assessment. The topics addressed include the use of the diploma or degree as a hiring requirement; the fairness of licensing and certification procedures as gateways to professions, with an emphasis on teacher certification; the use of IQ tests and other measures of general ability for classifying students; minimal competence assessment and the accountability of educational institutions; and the use of graduate school admission tests to select applicants fairly for the limited number of available places. The chapter concludes with a summary of the conditions under which the courts will tend to scrutinize competence assessment practices as well as suggestions for further legal research. (Other sections of the report are available separately--see note. The first is an overview; the last is a synthesis of issues.) (LRA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED192169

THE ASSESSMENT OF OCCUPATIONAL COMPETENCE

5. COMPETENCE ASSESSMENT AND THE COURTS: AN OVERVIEW OF THE STATE OF THE LAW

Arthur R. Block
and
Michael A. Rebell

REBELL & KRIEGER
New York, New York

Prepared for
The National Institute of Education
Contract NIE 400-78-0028

by
McBer and Company

February, 1980

CE027164

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

McBER 137 NEWBURY STREET · BOSTON · MASSACHUSETTS 02116

DEC 4 1980

TABLE OF CONTENTS

I. STATE OF THE LAW OVERVIEW	5.1
State of the Law Prior to Griggs	5.3
The Impact of Griggs	5.7
Washington v. Davis: A Dramatic Jolt	5.16
Developments Post-Davis	5.22
II. IMPLICATIONS FOR EDUCATION	5.33
Diplomas	5.37
Licensing, Certification, and Competency-Based Education	5.45
CBE and the NTE	5.50
Performance Evaluation	5.58
I.Q. Tests and Ability Tracking	5.64
Student Competency Assessment	5.74
Educational Malpractice	5.75
Minimum Competency Testing	5.79
Graduate School Admissions Tests	5.84
III. CONCLUSION	5.94
APPENDIX: The Emerging "Middle-Ground" Equal Protection Test and its Application to Education Law Litigation	A5.1
FOOTNOTES	

I. STATE OF THE LAW OVERVIEW

Over the past decade, the federal courts have become increasingly active in mandating compliance with specific competency assessment standards. In dozens of cases, they have invalidated test instruments, education requirements and experience requirements that private employees and public agencies had used in their hiring and promotion decisions. In response to these trends, some commentators have claimed that the courts "...are imposing a standard of excellence that does not exist in other areas of decision making."[1] Others have predicted that insistence on rigorous validation standards will totally discourage the use of objective employment tests, and substitute in their place either overly subjective practices, or else hiring based on racial quotas.[2]

The empirical findings reported in the earlier chapters of this study tend to substantiate the view that the validation standards being promulgated by the courts are far more rigorous than the current state of the practice.[3] On the other hand, these court rulings have taken a general direction that is consistent with, and often directly incorporates, the consensus of informed professional opinion among psychological and psychometric experts on these issues.[4] In fact, on certain issues, where controversy existed concerning the feasibility of fully

complying with certain professionally established standards, the courts have tended to take a moderating approach.[5] Indeed, the trend of the most recent court decisions is toward increased "flexibility" (but continued enforcement of substantive standards) in response both to a softening of administrative regulations and to a better understanding, learned from experience in previous cases, of the practical realities of implementing better assessment methods.

The substantial gap between the current state of the practice in competency assessment and existing regulatory and judicial enforcement standards raises major public policy questions concerning the informational basis upon which such standards are promulgated, and the extent to which laws and regulations can--or should--attempt radical revisions in well-entrenched employer practices. The implications of the discrepancy between the ideal and the actual, and barriers to improving practice are two of the themes that will be directly addressed in the concluding chapter of this study.[6] The present chapter will lay a groundwork for discussion, in addition to performing the more basic function of informing those in the field of the specific standards currently being enforced by the courts, and the likely directions for future rulings.

Much of the attention of psychologists and educators concerned with legal standards for competency assessment has focused on well-publicized Supreme Court decisions to the

neglect of important developments in the federal trial courts and lower appellate courts. These court rulings, however, are the main source of the specific mandates that public and private employers are expected to follow on a day to day basis.[7] They also constitute a sine qua non for explaining the context and significance of the major Supreme Court rulings. Consequently, the discussion which follows will analyze the decisions of the federal courts at all levels.

The State of the Law Prior to Griggs

In Title VII of the Civil Rights Act of 1964, Congress enacted a series of provisions broadly prohibiting employers from engaging in discriminatory practices in their hiring procedures.* Congress was aware that the use of employment tests was on the rise, and that in some quarters, unreliable and biased tests were being used in a manner that deprived minority groups of fair employment opportunities.[8] For these reasons, Congress intended to include employment testing practices and other competency assessment practices within the structures of the act. However, because of strong feelings raised by a decision of a hearing examiner for the Illinois

*42 U.S.C. S2000e-2. Originally, employees of private educational institutions, and of state and local government agencies were exempted from the Act; in 1972, however, most of these exemptions were abolished and employees of state and local governments were henceforth to be covered by the anti-discrimination mandates of the law. (42 U.S.C. S2000e-1).

Fair Employment Commission which suggested that standardized tests on which whites performed better than blacks could never be used,[9] a provision was incorporated in Title VII which specified that it shall not be an unlawful employment practice for an employer "to give and act upon the results of any professionally developed ability test" provided that such test is not "designed, intended or used to inate because of race..."[10]

Title VII also established the Equal Employment Opportunity Commission, and entrusted the EEOC with powers to investigate allegations of discriminatory employment practices and to seek conciliation agreements to rectify any discrimination found.[11] In its early attempts during the mid-1960s to effect compliance with the Act, the EEOC apparently determined that conducting ad hoc negotiations with individual companies was not proving fully effective: Consequently, the Commission decided to issue broad guidelines promulgating standards that would be used uniformly in its compliance endeavors.[12] Although the Commission had the authority to issue such guidelines pursuant to 42 U.S.C. S2000e-12, such guidelines, not being formal rules or regulations enacted in accordance with procedures under the Administrative Procedure Act, normally would constitute "interpretive material," indicating the agency's interpretation of the statutory intent. Guidelines of this type normally are held to be entitled to no greater weight than "other well-founded testimony by experts in the field of employment testing."[13]

A review of the major early Title VII cases indicates that, to the extent that their relevance was noted or accepted by the

courts, the EEOC Guidelines were used precisely in this manner. Thus, in United States v. H. L. Porter Company[14] and Colbert v. H-K Corporation,[15] the courts upheld the use of general aptitude and mental ability tests which clearly had not been validated in accordance with the EEOC guidelines.[16] After weighing the testimony of expert psychologists for the parties, the courts accepted a "common sense" standard, indicating that mental ability tests obviously had some relationship to the jobs at issue. As the court specifically stated in Porter, validation by a professional psychologist would not be required where a nonprofessional company personnel director, who appeared to be familiar with the job duties at issue, offered a "credible" opinion of the test's relevance to the job.[17]

Although in applying such a "common sense" standard, the courts, in the two cases cited above, upheld use of the general ability tests. In a greater number of these early cases, the courts invalidated testing practices under similar general "reasonableness" approaches. Thus, in Dobbins v. Local 212, International Brotherhood of Electrical Workers,[18] a test on electrical theory, which was apparently unrelated to practice on the job, was invalidated; in Arrington v. Massachusetts Bay Transportation Authority,[19] a general aptitude battery for bus drivers and toll collectors was held to have little business relevance; and in Hicks v. Crown Zellerbach Corporation,[20] use of the Wonderlic and Bennet Mechanical Comprehension batteries for production and maintenance employees was considered

unrelated to any valid business necessity.[21] But in only one of these cases (Hicks) did the court's decision focus on the relevance of the EEOC Guidelines in any detail; and in that case, the court specifically indicated that the Guidelines were consistent with the professional testimony in the case.

In short, although the courts were inclined to invalidate tests which were not job-related, especially in situations of gross adverse impact on minority job applicants, or of actual discriminatory intent, these decisions were made on a case-by-case basis, rather than by consistent applications of the EEOC Guidelines. The results emerged more from the judges' "common sense" impressions of the credibility of psychological testimony presented to them, and an independent reading of the precedential impact of prior cases, than from any general tendency to view the requirements in the EEOC Guidelines as being binding or entitled to great deference by the courts. In most of the cases where employment tests were invalidated, there had been no attempt whatsoever to undertake validation studies. As indicated by the decisions in H. K. Porter and Colbert, if any minimal attempt to show some degree of job-relatedness of these exams had been undertaken, the courts may well have accepted defendant's position, despite lack of professional validation in accordance with the EEOC guidelines.

The Impact of Griggs

The era of "common sense" judicial interpretation of the reasonableness of employment selection tests by the lower federal courts reached its culmination in the Griggs v. Duke Power Company cases. Prior to the effective date of the Civil Rights Act of 1964, the Duke Power Company had openly discriminated against black employees and job applicants on the basis of race. When, in response to the new law, the company abandoned its policy of segregating Negroes into the "labor department," it maintained two requirements for entry (by initial hiring or by transfer from "labor") into the other, more desirable departments: a high school diploma and successful completion of two professionally prepared aptitude tests.[22] At the time the case was filed, all but one of the 14 black employees of the company remained in the labor department, largely because of these barriers.

The United States District Court for the Middle District of North Carolina upheld the legality of these diploma and test requirements, and specifically rejected the EEOC's position that Title VII required a showing that the tests or diploma requirements at issue were specifically related to the duties of a particular job.[23] The court reasoned that 42 U.S.C. S2000e-2 required merely that employment tests be "professionally developed" and that the tests at issue (the Wonderlic Personnel Test and the Bennet Mechanical Comprehension Test) were well-known

standardized tests in the area, developed by competent professionals. In affirming the lower Court's decision, the United States Court of Appeals for the Fourth Circuit further indicated that the company had a valid business purpose in seeking to upgrade the general level of intelligence of its employees in a manner that would facilitate later internal promotions. The court specifically ruled that the EEOC Guidelines were not binding on the court because they were contrary to the legislative history which, the court held, indicated that Congress expected that general intelligence tests such as the Wonderlic would be permitted under S2000e-2.[24]

The central issue in the case when it reached the Supreme Court, then, was whether S2000e-2 should be interpreted (as held by the lower courts) to permit diploma requirements and the use of general intelligence tests "designed by professionals," or whether (as set forth in the EEOC Guidelines) a diploma requirement or an employment test would be considered an arbitrary barrier and not "professionally developed" if an adequate showing had not been made of its direct relationship to the duties of the job at issue. In Griggs[25], the Supreme Court unequivocally endorsed the EEOC's interpretation of the statute on these critical points:

> From the sum of the legislative history relevant in this case, the conclusion is inescapable that the EEOC's construction of S703(h) to require that employent tests be job-related comports with congressional intent.
>
> Nothing in the Act precludes the use of testing or measuring procedures; obviously they are useful. What Congress has

> forbidden is giving these devices and mechanisms controling force unless they are demonstrably a reasonable measure of job performance.... What Congress has commanded is that any tests used must measure the person for the job and not the person in the abstract.[26]

The crux of the Griggs decision* was expressed in the Supreme Court's statement that both a high school diploma requirement and an employment selection test must be "shown to bear a demonstrable relationship to successful performance of the jobs for which it was used."[27] However, since the Duke Power Company admitted that it had not undertaken any meaningful study of job-relatedness, the Court did not have occasion to spell out in detail the manner in which such a "demonstrable relationship" should be established.

Given the fact that the EEOC's interpretation of the critical issue of job-relatedness had been fully endorsed by a unanimous Supreme Court in Griggs, together with Chief Justice Burger's statement in his opinion that "the administrative interpretation of the Act by the enforcing agency is entitled to great deference,"[28] it was logical for the lower federal courts to assume that the specific, detailed validation requirements of the EEOC Guidelines henceforth should be considered the proper benchmark for analysis of Title VII cases. In contrast to their prior limited authority, the Guidelines were now given

---

*Griggs is also important, of course, for its holding that under Title VII an employment practice which has the effect of disproportionately excluding minority applicants will place a burden of justification upon the defendants, regardless of whether any actual intent to discriminate had been established.

virtually the binding effect of law. The predominant position of the lower federal courts was articulated by the Fifth Circuit Court of Appeals as follows:

> Their guidance value is such that we hold they should be followed absent the showing that some cogent reason exists for noncompliance.[29]

In the years following Griggs, the general pattern of judicial analysis in Title VII cases (after having found adverse impact on protected minority groups) was to closely analyze job-relatedness requirements, with specific reference to the EEOC standards applicable to the given situation, and often to measure the employment selection devices at issue in relation to those standards.[30]

This trend toward incorporation of the specific EEOC Guideline standards culminated in the Supreme Court's 1975 decision in Albermarle Paper Company v. Moody.[31] As the Court stated at the beginning of its opinion, one of the critical issues it sought to resolve in Albemarle was the precise question left open by Griggs as to the specific showing an employer must make "to establish that pre-employment tests...are sufficiently 'job-related' to survive challenge under Title VII."[32] Unlike the defendants in Griggs, the Albermarle Company had hired an industrial psychologist to study the job-relatedness of the two general ability tests, the Beta Examination and the Wonderlic, utilized by the company for hiring for jobs in a number of functional departments, each having one or more distinct lines of

progression. Albemarle's psychologist undertook a "concurrent validation" study dealing with ten job groupings selected near the top of nine of the lines of progression. Within each job grouping, the study compared the test scores of each employee with independent rankings of the employee, relative to each of his co-workers, made by two of the employee's supervisors. The results of this study indicated significant correlations in three job groupings for the Beta test, seven for the Wonderlic and two for both.

To a layperson unfamiliar with the technical requirements of test validation measurement, such a study may well have appeared to be reasonably job-related on a "common sense" basis. Such, in fact, had been the holding of the district court judge in this case. The Supreme Court, however, held that the fundamental benchmark for assessing compliance with Title VII job-relatedness requirements was to be the EEOC Guidelines which "draw upon and make reference to professional standards of test validation established by the American Psychological Association."[33] The Court then held that "measured against the Guidelines, Albemarle's validation study is materially defective in several respects."[34] Specifically, the Court pointed out that the validation correlations, obtained for certain job groupings, did not apply to each of the company's lines of progressions. The study did not meet the requirements of S1607.4(c) (2) of the Guidelines, because there was no evidence of "significant differences" between the categories of jobs

showing correlations and those which did not. Similarly, use of vague supervisory rating standards as a criterion measure were held to be in violation of S1607.5(b) (3), (4). In all, the court explicitly endorsed various provisions of the EEOC Guidelines, not fewer than eight times in the course of its opinion.*

The forcefulness of the court majority's incorporation of the EEOC Guidelines in its opinion was illustrated by the fact that two of the Justices (one of whom had been the author of the Griggs decision), felt compelled to register a protest against "the Court's apparent view that absolute compliance with the EEOC Guidelines is a sine qua non of pre-employment test validation."[35] Although acknowledging that the Guidelines are entitled to "great deference," Chief Justice Burger, in dissent, also protested against the majority's "wooden application of EEOC Guidelines" and reiterated the point that not being rules or regulations entitled to the force of law, the Guidelines should be given no more weight than other well-founded testimony by experts in the field.

Despite the reservations of the Chief Justice and Justice Blackmun, the other six Justices of the Court who held that Albermarle's validation study must be "measured against the Guidelines," decided that the EEOC Guidelines should--in effect,

*The Court's strong reliance on the Guidelines is further evidenced by the fact that nowhere in its validation analysis did the Court cite prior Title VII decisions of the lower federal courts.

if not in technical parlance--be given the full weight of law. Presumably, their decision was based on the view that in cases like Albemarle, involving a history of purposeful racial discrimination, the equal employment opportunity laws must be vigorously enforced and the Guidelines provided strong standards for doing so. In addition, given the increasing magnitude and complexity of test validation cases then winding their way through the courts, the judges undoubtedly felt that considerations of judicial economy made utilization of a consistent set of enforcement standards, promulgated by the federal agency having greatest expertise in the field, highly advisable.[37] The practical effect of the courts' adherence to the EEOC Guidelines following the Griggs-Albermarle decisions is dramatically illustrated by the overwhelming rate of plaintiff victories in the major Title VII litigations during this period. Our research has identified a total of 70 reported Title VII cases decided by the federal district courts and courts of appeals from 1971-76.[38] Of this total, plaintiffs were victorious in 56 cases, defendants in 13, and in one case involving dual validation issues, each side partially prevailed. Thus, plaintiffs won 80% of these litigations, a substantially greater proportion of victories than apparently was the trend prior to the Griggs decision.*

*In the subset of cases involving claims against public employers separately considered, we find that there was a total of 31 such cases and that plaintiffs were victorious in 26 of them or 84% of the time.

It is interesting to note that of the 14 cases where defendants prevailed, in whole or in part on test validation issues, in only one instance (Allen v. City of Mobile[39]) did a court base its ruling essentially on a "common sense rationality" analysis inconsistent with the EEOC standards. In six of the defendant victory cases, the courts held after analysis that defendants' practices were in compliance with EEOC requirements.[40] In the other cases, the courts did not find the EEOC requirements had been met, but held that for a variety of reasons Title VII job-relatedness provisions (the threshold issue for invocation of the Guidelines) need not be applied.[41]

In the large majority of cases where defendants' testing practices were invalidated, the courts steeped themselves in psychometric concepts and issued detailed opinions (often running 30 or 40 pages or more in length), which dissected the challenged practices in terms of the technical requirements for content validation, predictive validation, concurrent validation, etc.[42] As the Court of Appeals for the Second Circuit stated in Vulcan Society of New York City Fire Department v. Civil Service Commission, "Cases like this one have led the courts deep into the jargon of psychological testing."[43] By and large, the judges appear to have proved themselves adept in comprehending and applying this technical jargon. Although lower court judges were occasionally accused of confusing the concepts of construct and content validation,[44] in general, defendants' appeals were based on claims of overly vigorous

application of the standards to the facts at issue, rather than on any allegations of misunderstanding or erroneous application of the psychometric concepts.[45]

Although the courts in general rigorously applied the EEOC Guidelines, it is important to note (especially in light of the findings on the state of the practice set forth in Chapters 1-4) that the courts generally were aware (presumably through the testimony of expert witnesses) of professional reservations on a number of the specific requirements in the Guidelines, and tended to avoid basing their decisions on these controversial items. Thus, for example, although +1607.5(a) of the Guidelines specifically required a showing of criterion validation except where it could be shown that criterion validation "is not feasible," the courts repeatedly finessed the question of the feasibility of criterion-related studies in the particular situation, and tended to operate on the working assumption that the primary focus would be on compliance with content validation standards.[46] Similarly, the courts also tended to avoid the second major controversial aspect of the EEOC Guidelines: their insistence upon proof of the unavailability of alternative testing procedures having a lesser adverse impact, even if the test at issue was shown to be job-related,[47] although in several cases involving particularly egregious patterns of discrimination the available alternative standard was actually imposed.[48]

Thus, by 1976, the state of the law on Title VII test validation issues seemed relatively well settled. Basic adherence

to the EEOC standards was the rule of thumb, although the courts, without frontally challenging or undermining the basic applicability of the EEOC Guidelines, tended to deftly avoid fully imposing several of the most controversial features.* The Supreme Court's largely unanticipated decision in June, 1976 in Washington v. Davis,[47] however, sparked a basic reappraisal of many of the operating assumptions in this area.

Washington v. Davis: A Dramatic Jolt

The Supreme Court's ruling in Washington v. Davis had two major dimensions: (1) it was a pronouncement that, contrary to prevalent understandings in the lower courts, Title VII job-relatedness standards should not be applied in cases brought under the general provisions of the equal protection clause of the fourteenth amendment to the Constitution (as contrasted with cases brought directly under the Title VII statute), unless a finding had been made of actual intent to discriminate on the part of the defendants;** and (2) it upheld the legality of a

*Some courts noted that although basic adherence to the EEOC standards was essential, they would not require defendants to "justify every selection device to a mathematic certainty." Boston NAACP v. Beecher, 504 F.2d 1017, 1022 (1st Cir. 1974).

**Under Title VII, as the court had held in Griggs, job-relatedness requirements would apply if a discriminatory impact of the defendant's practices had been shown, regardless of the existence of any discriminatory intent. Since the 1972 amendments to Title VII had extended coverage of the Act to state and local goverment agencies, almost all employment discrimination cases in recent years have been brought under Title VII. (The Davis case had been filed prior to the 1972 amendments.)

civil service verbal ability test which clearly had not been validated in accordance with the EEOC standards.

Some commentators have interpreted this latter holding as being a substantial reversal of the Court's position in Griggs and Albermarle.[51] However, an analysis of the specific facts and setting of the case, as well as the Court's tortuous attempts to avoid directly confronting the issue of the relationship between this holding and its prior decisions, indicates that the Court's actual message in Davis, on this point, cannot be understood in such simple terms. Subsequent signals from the Supreme Court and post-Davis decisions of the lower federal courts indicate that the Griggs-Albemarle holdings remain viable; but since Davis, their application is undertaken in a somewhat more 'flexible" manner. In order to understand these developments, it is necessary to first briefly discuss the specific facts and rulings in the Davis case.

The employment selection instrument at issue was the personnel examination given throughout the federal Civil Service ("Test 21") which focused on verbal ability, vocabulary, reading, and comprehension. The defendant in the case, the Washington DC Metropolitan Police Force, had made significant strides in recent years to actively recruit minority applicants for police

---

Under these circumstances, the Court's likely objective in rendering this holding on the issue of intent was to avoid establishing a constitutional precedent with far-reaching implications for many other areas of government activity; the court probably did not mean to limit the applicability of the job-relatedness requirements. See 426 U.S. at 241.

jobs, and had been highly successful in this endeavor. A substantial number of blacks had been recruited and appointed to the police force, and, as the district court specifically held, the police department practices constituted "a model nationwide" of affimative action procedures. Despite this notable success, statistics indicated that blacks, to some degree, still failed Test 21 in numbers disproportionate to that of whites.* Although there were no validation studies showing whether success on Test 21 correlated with actual success on the job, the defendant's validation studies indicated a positive relationship between success on Test 21 and performance in the police training program. The district court had accepted these indications of "training course validation"[52] as sufficient proof of job-relatedness.

The Court of Appeals for the District of Columbia Circuit, although noting the department's impressive affirmative action attempts, held that the department must prove job-relatedness more directly. Success in a training program was an inappropriate criterion measure, absent proof that the curriculum of the program was related to the job. This selection procedure appeared to be primarily a measure of an ability to learn classroom materials, without being an actual predictor of a police officer's performance on the job. The appeals court reversed

*This adverse impact may, in fact, have been related to the very success of the affirmative action program, since active recruitment may have attracted a disproportionate number of less qualified black applicants. See Lerner, supra note 2 at 268.

the district court's holding,[53] noting that all prior cases which had considered this issue had uniformly rejected the concept of training course validation.

The Supreme Court's ruling on this issue reinstated the holding of the district court. The decision did not address the Title VII precedents concerning test validation upon which the court of appeals had ruled. The Supreme Court was able to avoid confronting those holdings because of the dichotomy it created in Davis between constitutional liability (based on intent) and statutory liability under Title VII (based on impact). Since Davis did not include a Title VII claim, and there was no basis for a claim of intentional discrimination (the department had a "model" affirmative action program) the constitutional claim was easily disposed of. The only relevant validation standards left to consider were regulations of the United States Civil Service Commission. These regulations, according to the Court majority's reading, specifically included "success in training" as a proper criterion for assessing the validity of a selection instrument.* Therefore, Test 21 was upheld.

---

*426 U.S. at 250. Justice Brennan, dissenting in Davis, claimed that the majority opinion misread the Civil Service Commission regulations on this point. 426 U.S. at 261-262. The Justice Department, citing "evolutions in the field" (such as the 1974 revisions of the APA Standards), had requested a remand for further consideration by the district court on this issue. Note also that the Civil Service Commission, as one of the federal agencies endorsing the 1976 Federal Executive Agency Guidelines and 1978 Uniform Guidelines, presently does not accept the use of a training course validation criterion which has not been shown to be related to actual performance on the job. See Uniform Guidelines SS14.B(3), 14.C(7).

Even though the Davis holding was not technically a decision applying Title VII, certain statements in the decision seem to suggest an interest in moderating somewhat the strict application of validation standards, even in Title VII situations. Although the Court went out of its way to base its training course validation holding on now outdated civil service regulations, which technically have no relationship to Title VII requirements, Justice White, writing for the majority also went on to hold:

> Nor is the conclusion foreclosed by either Griggs or Albemarle Paper Co. v. Moody, 422 U.S. 405 (1975); and it seems to us the much more sensible construction of the job relatedness requirement.[54]

This language could mean that in any test validation case, even under Title VII, if the defendants have, in good faith, implemented an effective affirmative action program,* and the test at issue involves a relatively simple verbal ability test which is "obviously" relevant to the job at issue, adherence to technical EEOC job-relatedness requirements may be relaxed.**

---

*This position is consistent with the "bottom line" concept in S4.C of the 1978 Uniform Guidelines which indicates that if the basic total effect of the employer's practices is non-discriminatory, enforcement will not be focused on possible minor, specific violations. See also Blumrosen, Developments in Equal Employment Opportunity Law - 1976 36 FED. B.J. 55 (1977), Furnco Construction Corp. v. Waters, 438 U.S. 567 (1978) (word of mouth hiring procedure resulting in high minority representation upheld).

** Note also Justice Stevens' comment in his concurring opinion that qualifications of law enforcement officials is an area "in which the federal district judges have greater expertise than in many others," and that Test 21 is a minimal verbal

In sum, the best interpretation of the overall implications of the Supreme Court's complex consideration of test validation issues in Davis would appear to be that the basic substance of Title VII job-relatedness standards, as articulated in the agency guidelines, should continue to be enforced; but that if a defendant is acting in good faith and a simple entry level examination has obvious job relevance, courts should not be overbearing in insisting on technical psychometric requirements. That the Supreme Court intended to promote continued, though perhaps more flexible, enforcement of the EEOC Guidelines, is indicated by positions taken by the Court in later cases involving enforcement of EEOC Guidelines. For example, in invalidating height and weight requirements which had a

---

ability test which appears "manifestly relevant to the police function." (426 U.S. at 254-255)

But cf. Note, "Developments in the Law, 1975 Term," 90 Harv. L. Rev. 56, 122 (1976): "The majority's emphasis on the importance of verbal skills in fulfilling the requirements of police work suggests that it may have found training course validation acceptable only because of an independent belief that the challenged test was indicative of skills necessary for successful job performance. The petitioner's brief cited in detail various Commission reports on police functions which would support such an impression. But the inherent unreliability of such untested beliefs is the reason that the EEOC regulations and courts have required objective empirical studies."

Note in this regard that, contrary to the opinion of some commentators such as Lerner, supra, note 2, it was the "complexity" of police work, as contrasted with the simple labor tasks at issue in Griggs, which promoted the Supreme Court's "flexible" approach. In fact, it appears that only because Test 21 was perceived as being "simple" and "obviously" related minimal job-related abilities, that the Court felt confident in upholding its use, despite the absence of rigorous validation evidence.

discriminatory impact on female applicants for correctional officer positions, in Dothard v. Rawlinson[55], the Court cited Griggs, Albemarle, the EEOC Guidelines, and Davis in its discussion of job-relatedness requirements, indicating a presumption of continued compatability of prior pronouncements with Davis. Even though a number of lower courts, even before Davis, had not applied specific job-related validation requirements in cases of this type, the Supreme Court now held that in order to establish the bona fides of the height/weight requirements, a test for applicants that directly measures strength should be adopted and validated in accordance with Title VII standards. Furthermore, in a number of other post-Davis cases, the Court specifically repeated its statements that EEOC Guidelines are entitled to great weight and cited Griggs and/or Albemarle as examples of proper application of EEOC Guidelines.[56]

Developments Post-Davis

For the period since the Supreme Court issued its decision in Davis, we have identified 42 reported test validation decisions by the lower federal courts.[57] Of this group, plaintiffs prevailed in 18 cases, defendants in 19, and 5 involved remands or other situations whose outcomes were as yet undetermined. In the decided cases, then, the plaintiffs have prevailed in only 47%.*

*Out of the subset of 24 public employer decisions in the post-Davis sample, plaintiffs have prevailed in 11 and defendants in 13. In these situations, then, plaintiffs prevailed 46% of the time.

These statistics, when compared with the much more striking pattern of plaintiff victories for the similarly selected sample of pre-Davis cases reported, supra at [p. 15], would appear to substantiate our expectation that Davis is being interpreted by many courts as calling for a "more flexible" approach to defendants' validation evidence. However, the entire statistical difference probably cannot be attributed to the impact of Davis. The very fact that Title VII and the Guideline standards have been in existence for almost a decade have led many to increasing sophistication and improvement in testing practices (or in the manner of presenting them to courts), at least by the large employers who might expect to have their hiring practices closely scrutinized by plaintiff groups or enforcement agencies.* Whereas defendants in the early cases often stipulated that no validation measures whatsoever had been undertaken, the issue in the "second generation" cases generally has been whether validation studies, which defendants did undertake, met legal requirements. Given these realities, the fact that plaintiffs continued to prevail in almost half the cases substantiates the view that, despite the trend toward "flexibility" heralded by

*Compare on this point the hypothesis in Chapter 1 of this study [p. 118] that mature organizations "...are most likely to support a program of selection validation...". Chapter 1 further hypothesizes that the prime strategy of an organization which is "visible" will be "to document its selection practices" (i.e., by showing a lack of adverse impact) rather than to validate its proved defendant success rate in major court cases. This would be consistent with either enhanced documentation (and increased sophistication in presentation of available evidence) or improved validation procedures.

Davis*, the courts continue to apply substantially the relevant validation standards.

A strong commitment to continued adherence to the requirements of the Guidelines has, in fact, been explicitly articulated by most of the lower courts. For example, the Court of Appeals for the Seventh Circuit in United States v. City of Chicago[58] specifically reiterated the Georgia Power standard that compliance with the guidelines will be required, absent a showing of cogent reasons for their non-use.[59] Also significant in this regard is the recent decision in Allen v. City of Mobile[60] in which Judge Pittman, the same judge who had refused to follow the Guidelines in 1971 (immediately after Griggs), relied on Albemarle and Georgia Power and held that the Guidelines would be rigorously applied (and, specifically in this case, the highly controversial requirement to consider alternative methods having less discriminatory impact). The decision in Allen, announced the same week the 1978 Uniform Guidelines went into effect, also indicated that the now

*This trend was also apparent in the promulgation, shortly after the Davis ruling, of the new Federal Executive Agency Guidelines, which modified the EEOC standards in many critical areas, such as the statistical definition of adverse impact, preference for criterion-related validation, insistence on empirical job analyses, requirements for differential validation, and searching out of selection methods with lesser impact. (See, e.g., 41 Fed. Reg. 51744 Nov. 23 1976) Because of these modifications, the EEOC refused to adopt the FEA Guidelines which were endorsed by the Justice Department, the Labor

unified position of the federal enforcement agencies would add additional clout to the Guidelines.[61]

Perhaps the most graphic way to illustrate the complex interplay (in a general context of improved defendant testing practices in the post-Davis era) between continued, although "flexible" enforcement of the Guidelines, is to compare the handling, in a similar setting, by the same judge of employment selection issues before and after the Davis decision. The two Bridgeport Guardians cases,* both decided by Judge Newman of the Connecticut District Court (the first in 1973, before Davis, and the second in 1977, after Davis), provide a rare opportunity to engage in such a comparative analysis.

Department, and the Civil Service Commission. Instead, the EEOC re-issued its original 1970 Guidelines. The existence of this inter-agency competition was, of course, quickly made known to the courts. Generally, instead of preferring one set of standards or the other, the judges would purport to apply both in their large areas of overlap. See, e.g., Friend v. Leidinger, 446 F. Supp. 361, 367 (E.D. Va. 1977); affd. 588 F.2nd 61 (4th Cir. 1978); Dickerson v. U.S. Steel, 17 E.P.D. 18528 (E.D. Pa. .978); NAACP Ensley Branch v. Seibels, 13 E.P.D. S111504 (N.D., Ala. 1977). In 1978, all these federal agencies finally agreed to jointly accept a new set of Uniform Guidelines (See n. 4, supra.) Most of the FEA modifications were maintained in the 1978 Uniform Guidelines, although on some points, such as specific requirements for content validation, the 1976 version moved closer to the original EEOC position.

*These cases, involving challenges to hiring practices of a municipal police force, are also obviously close in their facts to the precise situation in Davis. An overview of those cases (which are quite numerous) in our sample involving police department hiring practices, provides some interesting insights. Of 15 such cases in the pre-Davis sample, plaintiffs were upheld on the validation issues in 12 instances, with courts often holding that general intelligence or basic skill tests, somewhat similar to Test 21, were invalid because they were not job-related. See e.g., Castro v. Beecher, 334 F. Supp. 930 (D. Mass

The first of those cases, Bridgeport Guardians v. Members of Bridgeport Civil Service Commission et al.,[62] (referred to as "Bridgeport I") concerned a challenge primarily to the use of a standardized ability test battery (covering reading comprehension, vocabulary, arithmetic, exercise of judgment, general information and "capacity to observe and remember faces and basic biographical data"),[63] which was used on both a cut-off score and rank list basis for selection of municipal police officers. After holding that there was a clear pattern of adverse impact (the passing rate for whites was more than 3 times the rate for blacks and Puerto Ricans), the court reviewed in some detail the "standards applied in the field of psychological testing."[64]

In doing so, Judge Newman displayed a sophisticated understanding of the psychometric techniques, noting, for example, the difficulty of establishing precise criterion measures for complex jobs. Citing the provisions of the EEOC Guidelines then in effect, which stated a clear preference for predictive

1971), aff'd, 459 F.2d 755 (1st Cir. 1972). Although as indicated in the main text, this predominant pattern of plaintiff victories seems to be changing, in the few post-Davis police cases reported to date, plaintiffs have still been successful in overturning police examinations. Apparently, the limited scope of Test 21 in Davis was not typical of the more extensive police entrance batteries used in other areas and the courts, therefore, continued to strike down these tests as not being job-related. See, e.g., U.S. v. City of Chicago, 599 F.2d 915 (7th Cir., 1977), cert. den., 934 U.S. 875 (1978); Guardians Association v. Civil Service Commission 431 F. Supp. 526 (S.D.N.Y. 1977), vacated and remanded, 562 F.2nd 38 (2nd Cir. 1977), on remand 466 F.Supp. 1273 (S.D.N.Y, 1979); NAACP, Ensley Branch v. Seibels 13 E.P.D 11, 504 (N.D. Ala. 1977), Allen v. City of Mobile 464 F. Supp. 433 (S.D. Ala. 1978) Scf. Detroit Police Officers v. Young, 466 F. Supp. 979 (E.D. Mich. 1978).

validation, the Judge indicated that a concurrent study probably could have been undertaken by the employer in this case, since test scores were still available for a substantial number of the present patrolmen and the suit had been pending for nearly a year. Nevertheless, since no attempt at such a concurrent validation study had been undertaken, the court turned to a consideration of defendant's assertions that content and construct validation had been adequately demonstrated. He determined that professional standards of content and construct validation had not been met because of the lack of an adequate job analysis:

> There is no evidence that the job of a Bridgeport patrolman has ever been analyzed to determine what knowledge an applicant should be required to possess or what constructs should be identified in a prospective patrolman. The current job descriptions are rudimentary at best and most likely inadequate for such analysis, had it been attempted. There has been no showing that the exam measures with proper relative emphasis all or even most of the essential areas of knowledge and the traits needed for proper job performance.[65]

Judge Newman's decision was subsequently affirmed by the United States Court of Appeals for the Second Circuit which noted that "Judge Newman meticulously reviewed the evidence." In its decision, the Second Circuit, as might be expected from an appellate court which did not hear direct testimony, emphasized the facial lack of content validity, with reference to specific examination questions, rather than the more complex defects in the process of undertaking an adequate job analysis elaborated by Judge Newman. The court also noted the arbitrariness of the cut-off score requirement, which apparently

had been mandated by a city ordinance, without reference to the requirements of the specific police job at issue.

In 1977, Judge Newman had occasion to rule on another challenge to examination procedures of the Bridgeport Police Department. In this later case, Bridgeport Guardians v. Bridgeport Police Department[67], (Bridgeport II) the dispute was about the promotion examination for police detectives.* In contrast to his 1973 ruling, Judge Newman now held that the exam appeared to satisfy applicable job-relatedness requirements and he refused to issue a preliminary injunction invalidating the use of the examination. A review of the facts in the latter case indicates that the differing outcome in Bridgeport II stemmed both from a marked improvement in the defendant's testing and validation procedures, and a post-Davis judicial inclination to apply the basic substance, but not each technical detail, of the EEOC Guidelines.

After Judge Newman's original ruling in Bridgeport I, city officials had hired McCann Associates, a management consultant firm with substantial experience in civil service personnel matters, to "prepare an exam free of any racial or cultural bias."[68] The firm undertook an extensive job analysis. Interviews were conducted with numerous police detectives and captains of the relevant divisions. The firm also reviewed

*Promotion procedures had also been challenged in the 1973 case, but because of a lack of evidence of discriminatory impact, the court did not review the job-relatedness of those procedures at that time.

statistics in the detective bureau files which provided indicia of the types and quantity of work done by local detectives. Based on this information, a job description was prepared which identified 20 categories of tasks, within each of which 10 to 20 specific tasks were set forth. This job description was then compared with a subject matter checklist which the firm apparently regularly utilized in its work on police examinations. Fourteen of the 52 areas on the checklist were selected as being pertinent to the duties of detectives in Bridgeport. Examination questions were then prepared reflecting the relevant weight and importance of each function in Bridgeport, as indicated by job analysis. As a final step, a reading index was applied to the entire examination to establish the degree of reading difficulty (the index revealed that the exam was at the eighth and ninth grade levels of reading difficulty).*

A comparison of the detailed job analysis undertaken in Bridgeport II, with the impressionistic job descriptions utilized in Bridgeport I, dramatically illustrates the substantial changes in practice that were implemented as a result of the

*The cut-off score of 75 on this examination was determined by the McCann firm based on the fact that the average national correct answer rate on these questions was 68%. Based on a determination that a somewhat higher level of proficiency was desired for Bridgeport detectives (and influenced by the Civil Service Commission rule mandating a 75% score on all civil service exams), an additional 7% was added to this national average. The court held that this cut-off approach was acceptable under S1607.6 of the EEOC Guidelines, which specified only that cut-off scores "will be reasonable and consistent with the normal expectations of proficiency within the work force...".

court's original intervention in this area. Judge Newman, in contrast to his sharp criticism of the lack of adequate job analysis in Bridgeport I, now stated that the steps outlined in the preparation of this job analysis and examination "disclose an adequate level of skill, objectivity, and concern with achieving content validity."[69] Plaintiffs' expert in the case was the "ubiquitous" Dr. Barrett,[70] who had testified on their behalf in the Bridgeport I. He criticized some of the steps taken in the job analysis and test preparation procedures because of their lack of strict conformity with the APA standards. However, the court, while noting that the APA standards are "relevant to an assessment of an exam challenged in a Title VII suit," nevertheless, summarily set aside Dr. Barrett's objections as amounting to technical deficiencies which did not substantially affect the fact that "the end product achieved sufficient content validity."*

*431 F. Supp. at 937. Dr. Barrett also challenged the basic content of the examination questions themselves. Using specific questions as examples, he claimed that they did not really distinguish between those who can adequately perform on the job and those who can not. For instance, he objected to questions that required abstract knowledge, unrelated to specific job tasks, such as a question concerning national statistics on the incidence of forgeries. Although exhibiting some degree of skepticism concerning the significance of these criticisms, Judge Newman, nevertheless, decided to fully consider the objections. Specifically, he undertook a "judicial examination reading" in order to reach a reasonable conclusion as to the frequency of the type of questions being criticized by Dr. Barrett and in order to detemine whether, assuming arguendo the validity of the criticisms, they substantially affected the overall outcome. Although, as the Judge himself acknowledged, judicial exam reading has its own dangers in the

In short, Judge Newman's approach to testing validation issues in Bridgeport II exemplifies the pattern in post-Davis cases of flexible but substantial enforcement of the EEOC Guidelines in a context of dramatically improved employer testing and validation practices.* Apparently, taking a cue from the Supreme Court's approach in Davis, Judge Newman carefully applied the Guidelines, but at the same time he gave short shrift to highly technical objections of plaintiff's expert witness,

---

context of police functions which "are not foreign to the judicial experience (431 F. Supp. at 937, cf. Washington v. Davis, supra n. 50 at 255 [Stevens, J., concurring]), he considered such an undertaking to be justified.

After personally analyzing each question on the exam, the Judge determined that there were eight questions for which arguably more than one of the multiple choice answers might be correct. In addition, he identified ten questions which, on their face, may have been unsuitable because the knowledge tested for seemed unrelated to actual performance of a detective's duties or was otherwise unsuitable. Rather than reaching the issue of whether a finding of eighteen possibly questionable items by a judge who does not claim to have psychometric skills should provide the basis for invalidating part or all of the examination, Judge Newman decided to ask the defendants to rescore the tests, eliminating the ten unsuitable questions and allowing credit for an additional correct answer to the eight multiple choice questions to see if the results would have been materially affected. The results of this rescoring indicated that they would not; there was a very slight variation in the ranking of candidates, but the rescoring did not produce a passing grade for any black candidate, and did so for only one Hispanic applicant.

*In contrast to the highly favorable affirmative action stance of the police defendants in Davis, in Bridgeport, despite the 1973 ruling in Bridgeport I, there still were no black or Hispanic supervisors on the police force. This fact, which was emphasized by Judge Newman, undoubtedly affected the degree of scrutiny which he applied to the challenged practices, which nevertheless were upheld.

once he was satisfied that there had been substantial compliance with basic professional and EEOC content validation standards. It is impossible to weight these two factors--improved validation procedures and moderated judicial review standards--but it appears that both contributed to the reversal of the police department's fortunes in the second case.[71]

II. IMPLICATIONS FOR EDUCATION

The legal standards developed in the Title VII employment selection context have important direct and indirect implications for education. Title VII, by its terms, directly applies to the hiring practices of local government agencies, including school boards. In addition, as the Supreme Court stated in Griggs, diplomas issued by secondary schools and higher education institutions are also subject to Title VII job-relatedness requirements, if utilized as hiring or promotional criteria by employers. Aside from these areas of direct applicability, the courts' extensive scrutiny of testing instruments under Title VII also provides an important precedential point of reference against which the validity and suitability of a wide variety of educational testing practices may be measured. For example, in recent cases and legal commentary concerning public school tracking assignments,[72] minimum competency testing,[73] and admission tests,[74] direct reference was made to Griggs and other employment testing cases.

The relationship between the court holdings in the employment testing cases and analogous situations in specific areas of educational testing is complex and cannot be described in precise, concrete terms. On the one hand, it cannot be assumed that every educational test brought before the courts (and

especially those not directly related to an employment selection practice) will be closely scrutinized under the specific EEOC Guideline standards, or under the particular precedential holdings of the Title VII cases. Indeed, Washington v. Davis made clear that the strict EEOC validation standards will not automatically be applied in constitutional challenges not brought directly under Title VII.* On the other hand, the very fact that during the past decade the courts have involved themselves in detailed analyses of assessment practices and have become educated in psychometric techniques is a reality that cannot be lightly dismissed. A contemporary court is not likely to dismiss challenges to school testing practices out of hand because of a presumption that courts are not capable of reviewing decisions of educational experts. Given this experience, courts are now more inclined to involve themselves in difficult areas of educational judgment, if strong facts calling for remedial action are presented in a particular case.

To date, there have been a number of significant judicial cases involving educational testing practices. These particular decisions, although not yet calling for broad, judicial

*As our brief discussion of the pre-Griggs cases indicated, until the direct applicability of Title VII and the specific EEOC guidelines was made clear, the courts were hesitant to undertake vigorous initiatives, even in the employment discrimination field. In other words, it was legislative and administrative action in passing Title VII and articulating specific guidelines that induced the courts to rigorously scrutinize employment testing practices. Without analogous specific statutory impetus in regard to educational testing practices, similar intensive scrutiny by the courts cannot be anticipated.

scrutiny of all educational testing situations, not only provide precedent for further development of constitutional principles as they affect the educational process, but also create a climate of "legitimacy" for reformers who seek to persuade legislative and administrative bodies to impose more stringent requirements on school testing practices. Thus, initial judicial forays, both into the rights of the handicapped and I.Q. testing-ability grouping practices, significantly influenced the passage of statutes setting forth detailed procedural and substantive rights for handicapped students,[75] and legislative and administrative prohibitions against unvalidated tracking practices.[76]

In short, there is a complex interrelationship between the judicial branch, and the legislative and executive branches in the fashioning and implementation of specific reform requirements in a newly developing area, such as the scrutiny of educational testing practices. The pattern of this complex interrelationship indicates that initial, tentative judicial pronouncements often provide legitimacy for further direct reform efforts of legislative and administrative bodies; new statutes and regulations which are enacted subsequently become the basis for further active involvement by the courts in enforcing new statutory oversight functions.[77] In this context, the history of substantial judicial and legislative activism in the employment testing area will no doubt have increasing impact on analogous testing situations in the

broader educational arena, although at the present time, the precise nature of this scrutiny cannot be forecast.

The state of the practice overview presented in Chapters 1-3 of this study indicates that the actual test validation practices of employers fall far short of the standards articulated by professional experts, the APA, and the regulatory agencies. Judicial enforcement appears to fall in mid-range area. Courts require a higher level of validation than appears to be the general practice, but especially where the good faith of the defendants has been established, the judges tend to avoid insisting on strict adherence to all technical professional-regulatory requirements. They employ a "flexible" approach. Assuming that this pattern will also be applicable to the broader educational testing domain, educators should increasingly expect to be held accountable for the validity of their testing practices; but such accountability, when enforced by the courts, will be tempered by a cognizance of practical realities.

From the perspective established in these introductory remarks, the pages which follow will consider the direct and indirect application of the legal precedents established in the Title VII employment testing validation cases to the following five specific areas involving testing in the educational context:

1. Diplomas;
2. Licensing, certification, and competency-based education;

3. I.Q. testing and ability-level tracking;
4. Student competency assessment measures; and
5. Tests for admission to higher educational institutions.*

Diplomas

In its landmark decision in Griggs v. Duke Power Co., the Supreme Court held that not only employment selection tests, but also the requirement that job applicants hold a high school diploma, must be shown to have "a demonstrable relationship to successful performance of the jobs for which it was used."[78] As with ability tests, the court indicated that employers would not be permitted to generally "upgrade" the calibre of their employees by demanding proof of successful completion of a high school curriculum, if it could not be shown that such diploma requirements (which had a substantial detrimental impact on minority applicants) were not directly relevant to the jobs at issue. The Court spoke in strong language on this point:

> History is filled with examples of men and women who rendered highly effective performance without the conven-cional badges of accomplishment, in terms of certificates, diplomas, or degrees. Diplomas and tests are useful serv-ants, but Congress has mandated the common sense proposition that they are not to become masters of reality.[79]

---

*A brief discussion of the emerging "middle ground" standard for constitutional equal protection analysis is presented in an appendix to this chapter. In the absence of specific statutory or regulatory mandates, this approach provides the most direct legal basis for court involvement in the latter three of the areas listed above.

The Supreme Court's statements in this regard stand in stark contrast to the clear trend in recent years for employers to require ever-increasing years of schooling from applicants for jobs of all types.[80] Indeed, efforts by large city school districts to stem rising high school drop-out figures often are based upon the argument that, without a high school diploma, one cannot expect to obtain employment, or desirable employment.

The immediate implications for education of the Griggs holding on this point would appear to be twofold. On the one hand, invalidation of diploma requirements, at least for lower level blue-collar jobs, might undermine current inducements for students who are not academically motivated to remain in school solely for the purpose of completing degree requirements.* On the other hand, the decision could serve as an impetus for improving both the quality of the education and the specificity of the assessment measures which the diploma purportedly reflects. In other words, if diplomas can be seen as accurate indicators of specific levels of competence which are relevant to job functions (whether at blue-collar or advanced managerial levels), their importance would become significantly enhanced.

---

*These tendencies would, of course, be most directly relevant for high school vocational programs and professional training courses at the higher education level. Decreased significance of the vocational relevance of educational diplomas might be viewed by some educators as a salutary return to emphasis upon the value of "pure" liberal arts programs. See Huff, "Credentialling by Tests or Degrees: Title VII of the Civil Rights Act and Griggs v. Duke Power Comany," 44 Harv. Ed. Rev. 296. 259-262 (1974).

To date, the court decisions in this area indicate that, despite the Supreme Court's dramatic language in Griggs, a flexible "common sense" approach to the issue is being undertaken. Although diploma requirements for lower-level workers (like the plaintiffs in Griggs) which do not have any apparent rational justification, tend to be peremptorily invalidated in other areas (especially those involving higher-level jobs calling for college degrees), the courts have been hesitant to lightly discount the significance of academic credentials. In short, the judicial approach to the diploma issue is comparable to the post-Davis "flexible" approach discussed in the first section of this paper. Interestingly, this more flexible approach was adopted by the courts from the outset in this area, and the decision in Washington v. Davis did not result in any dramatic shift in approaches taken by the lower federal courts.

The series of post-Griggs diploma decisions did not reflect the same degree of deference to the EEOC Guidelines that was evident in cases involving ability tests. The EEOC, having taken the position that a diploma requirement is a "test,"[81] indicated that they must be subject to the same strict validation standards as are ability or achievement tests. The Supreme Court in Griggs, however, did not clearly endorse the EEOC's position on this point. Although the Court obviously assumed that diplomas were among the employment selection devices subject to the requirements of Title VII, its specific endorsement of the EEOC Guidelines in that case was limited to

the agency's interpretation of the phrase, "professionally developed ability tests," under Section 703(h) of the Act.* Just prior to its strong enforcement of the EEOC guidelines on that issue, the Court stated:

> Section 703(h) applies only to tests. It has no applicability to the high school diploma requirement.[82]

Thus, although the lower courts interpreted Griggs as endorsing the specific, rigorous EEOC Guideline requirements as they applied to employment testing, judicial opinion was mixed as to whether the brief comment in footnote 8 was meant to limit the endorsement of specific EEOC positions to ability tests, or was merely a technical, passing commentary which was not meant to undermine a general endorsement of EEOC Guidelines for all selection devices.[83]

The question as to whether the EEOC Guidelines should be directly applied to diplomas clearly is not an abstract, technical issue. Rigorous, empirical EEOC requirements would be nearly impossible for employers to meet in most diploma situations. For example, in terms of content validation, the diversity of curriculums and diploma standards in the thousands of schools throughout the country means that there are no uniform, readily identifiable diploma "contents" against which

---

*See discussion at pp. 9-10 supra.

particular job descriptions can be matched.*

In order to obtain an overview of the specific judicial activity in this area, we surveyed the period 1968 through October 1979 for all reported cases that rendered liability decisions on the issue of whether a diploma requirement was job-related or a business necessity.[84] This survey identified 33 relevant decisions.[85] Of these, two cases had been decided pre-Griggs, and in both these instances, the lower courts had assumed with little or no discussion that a high school diploma requirement was acceptable.[86] After Griggs, the courts invalidated high school diploma requirements in 19 cases, and upheld them in only five. Interestingly, all five of the cases in which defendants prevailed involved high school diploma requirements for police jobs. In addition, of seven cases involving higher education requirements, defendants won five. Thus, the overview of the survey of cases indicates a striking pattern. High school diploma requirements for lower-level jobs

*If this issue were pressed, courts might conceivably accept expert educational testimony concerning the typical content of a high school or college curriculum in the educational or occupational field at issue, or perhaps in the particular geographic locale from which the employee applicant pool is derived. (See discussion of United States v. South Carolina, infra at p. 56. ff.). In one diploma case, the plaintiffs used expert testimony to prove an absence of content validity. The educational expert there testified that the mathematical skills which the defendants alleged were necessary for satisfactory performance in an apprenticeship program were taught in the fourth through the ninth grades of most school systems. Therefore, a high school diploma requirement would exclude many qualified persons. EEOC v. Local 638, 401 F. Supp. 467 (S.D.N.Y. 1975).

were uniformly struck down, except in the police employment context, whereas higher education degree requirements, presumably related to higher level managerial positions, consistently were upheld.

The main factors that judges seem to have taken into account in these diploma cases were: (1) their perceptions of the defendants' good faith; (2) the facial plausibility of the relationship between the diploma requirement and the nature of the job at issue; and (3) the potential effect of striking down the requirement on important public health and safety considerations.

The 19 post-Griggs cases in which high school diploma requirements were invalidated involved blue-collar jobs: factory labor,[87] skilled trades,[88] or firefighter positions.[89] The geographical and historical context of these decisions appears to be significant. Ten of the 19 cases were set in the deep South, and were brought against employers who, prior to the passage of the 1964 Civil Rights Act, had apparently utilized intentionally discriminatory employment selection devices. The substitution of new diploma requirements for the previous overt discriminatory practices obviously raised substantial suspicions about defendants' good faith, when it was shown that the diploma requirements disproportionately affected black applicants.

Within this context of past discriminatory motives, the courts' scrutiny of high school diploma requirements was demanding, but nevertheless, the judicial analyses were generally

non-technical. The courts often struck down the defendants' diploma requirements, either by finding that the requirement on its face was not job-related, or by emphasizing the defendants' failure to even attempt to produce any evidence of empirical validation. That is, a court sometimes found, after comparing the duties of a blue-collar job with the content of a typical high school course of study, that large areas of the curriculum were irrelevant to job performance. For example, in U.S. v. Georgia Power Company,[90] the defendant company attempted to justify the diploma requirement on the ground that it provided a useful measure of reading comprehension ability. The court rejected this argument, however, stating:

> Many high school courses needed for a diploma (history, literature, physical education, etc.) are not necessary for these abilities. A new reading and comprehension test... might legitimately be used for the job need.[91]

A related ground for invalidating the diploma requirement was through application of the "business necessity test," which holds that an employer may not use a job-related qualification that results in discriminatory impact if there is an alternative, job-related selection device available which would not have such an effect.[92] Although questions about feasible alternatives may become extremely complex in regard to high level managerial jobs, available alternatives to these high school diploma requirements were apparent. Typically, since the diploma was justified as a proxy for some easily measurable skill such as reading comprehension, the courts could, with

little difficulty, assert that the defendants' interests could be easily satisfied by replacing the broad diploma requirement with a simple reading test.[93]

As indicated above, the only cases in our survey in which high school diploma requirements were upheld involved police officer positions.[94] While one might attempt to distinguish them from many of the other high school diploma cases by the fact that they were, with one exception, brought in northern cities where there was no history of overt or de jure discrimination, and the jobs at issue involved public safety functions, these explanations are insufficient. In four other cases involving firefighter positions in northern cities, the diploma requirements were struck down.* Moreover, since evidence of verbal communication skills was the main rationale offered for the diploma requirement in the police cases, it might have been appropriate to apply the "business necessity" doctrine, and call for the substitution of simple verbal ability tests.

The critical factor in the courts' acceptance of the diploma requirement for police jobs appeared to be a judicial perception that, on its face, a diploma requirement seemed related to police duties. In this regard, the courts relied heavily on the recommendations of independent commissions which

*Indeed, two of the five police situations were in cases where the court simultaneously invalidated the use of a diploma requirement for firefighter jobs. U.S. v. City of Buffalo, 457 F. Supp. 612 (W.D.N.Y. 1978); League of United Latin American Citizens v. City of Santa Anna, 410 F. Supp. 73 (C.D. Cal. 1976).

had considered high school diplomas to be job-related (even though the commission findings were not based on the kind of systematic empirical evidence required by EEOC Guidelines).[95]

As the court stated in City of Buffalo, supra:

> In the patrolman case, I find that the requirement has been validated by a "meaningful study of their relationship to job performance ability" as required by Griggs. I refer to the 1967 report of the President's Commission on Law Enforcement and Administration of Justice, "The Challenge of Crime In A Free Society," at 106-110, its underlying "Task Force Report: the Police," at 126-128, and the subsequent 1968 "Report of the National Advisory Commission on Civil Disorders" at 166. The level of education required for police officers was a central concern of these reports. The conclusion which they reached is that a high school education is a bare minimum requirement for successful performance of the policeman's responsibilities. This reasoning has been followed by several courts to uphold the high school requirement. [Citing Castro, Ballard, and League.] I find the reports and the cases cited to be ample support for continuation of the diploma requirement for police patrolmen.[96]

The higher education degree cases also reflected the pattern of "flexible" judicial reactions, avoiding technical EEOC requirements. In the two cases won by plaintiffs, employers with histories of overt, intentional discrimination against black employees had instituted college degree requirements for low-level administrative and supervisory positions.[97] The other five cases, won by defendants, dealt with qualifications for positions as a commercial airline pilot, university professor, campus security officer, public health program representative, and state narcotics agent.

Public health and safety seemed a primary concern of the court in finding for the defendant in Spurlock v. United Airlines,[98] the airline pilot case.[99] At issue was a

requirement that an applicant to airline pilot training school have a college degree as a prerequisite to taking the training admissions test. Although the defendant produced no empirical validation of the diploma requirement, the court accepted general testimony concerning the importance of selecting "the best qualified applicants" for this critical public safety position:

> When a job requires a small amount of skill and training and the consequences of hiring an unqualified applicant are insignificant, the court should examine closely any pre-employment standard or criteria which discriminates against minorities. In such a case, the employer should have a heavy burden to demonstrate to the court's satisfaction that its employment criteria are job-related. On the other hand, when the job clearly requires a high degree of skill, and the economic and human risks involved in hiring an unqualified applicant are great, the employer bears a correspondingly lighter burden to show that his employment criteria are job-related. Cf. 29 CFR S1607.5 (c) (2) (iii). The job of airline flight officer is clearly such a job.[100]

The court's decision in Scott v. University of Delaware,[101] was based on a "common sense" acceptance of the facial validity of a requirement that a university professor hold a Ph.D. or its equivalent, despite plaintiff's argument that this type of degree requirement had not been empirically validated in accordance with the EEOC Guidelines.[102] The court found, however, that the university's contention that the doctorate degree was meaningfully related to a person's ability to conduct research, think creatively, and add to the existing fund of knowledge through publication, was convincing on its face.*

*Compare with the campus patrolman decision, Jackson v. Curators of the University of Missouri, 456 F. Supp. 879 (E.D. Mo. 1978), which was based on the court's rather uncritical

The court concluded:

> While plaintiff is critical of the emphasis currently being placed on scholarly research and publication by the University, the University's choice of mission is not a subject for judicial review.[103]*

In short, then, although courts have invalidated diploma requirements in most cases where issue was joined on this point, the invalidation usually occurred in a setting of discriminatory motives or clear-cut arbitrariness in imposing diploma requirements for blue-collar positions which had no relationship to the job at issue. In cases involving diploma requirements for police officers and with higher education degree requirements, the court would act in a "flexible" manner and not vigorously enforce technical validation concepts.[104] In general, the courts have not involved themselves in scrutinizing the actual content of vocational or professional training programs for which the diplomas have been awarded.

---

acceptance of a police chief's testimony that two years of college education was needed by a campus patrolman because of the need to perform hazardous duties without supervision.

*The court also appeared to be impressed by the University's good faith in regard to the racial implications of its hiring practices. Overall, 27% of the university's faculty lacked Ph.D. degrees, but the percentage for the 12 black full-time faculty members was 50%. 455 F. Supp. at 1126 n. 66.

Licensing, Certification and Competency-Based Education

The analysis of the state of the art in Chapter 3 of this study indicated that current licensing and certification practices generally fall even further short of meeting professional validation standards than do industrial job selection techniques. The reasons for this pattern are complex and appear to include the political considerations that motivate many certification decisions, as well as the lack of an explicit profit incentive to improve test validity. An additional factor may be that the jurisdiction of the courts to scrutinize licensing and certification practices has not been clearly established and relatively little judicial enforcement of licensing standards has occurred to date.

Both the original EEOC Guidelines and the current Uniform Guidelines broadly interpret Title VII to apply to "selection procedures," defined as "any measure, combination of measures, or procedure used as a basis for any employment decision."[105] The enforcement agencies have assumed that the standards and procedures used by licensing and certification boards, at least to the extent that they have disproportionate impact on the employment of minority groups, should be subject to detailed scrutiny, under the guidelines.[106] However, a number of courts, especially in cases challenging the validity of bar examinations, have held that the definition of an "employer" under Title VII does not include licensing and certification boards.[107] Other courts have taken a contrary position.[108]

Meanwhile, both the EEOC and the Justice Department have continued to insist that Title VII should be interpreted to apply directly to licensing and certification boards.

Interestingly, however, certification of pedagogical employees has tended to constitute an exception to this general non-enforcement trend, apparently because the courts' jurisdiction to scrutinize the hiring practices of local school districts could clearly be established without reaching the still unresolved issue of whether the state certification boards, whose mandates applied by the districts are themselves directly subject to Title VII requirements.[109] In other words, if a school district can only hire state certified applicants, state certification procedures become subject to scrutiny as an employment selection device, rather than as a general licensing requirement. Much of the judicial activity in this area has concentrated on the reliance of local school districts, mainly in the South, on the National Teachers Examination (NTE) as a certification standard. These cases put into direct focus the fundamental issue raised by the competency-based evaluation movement discussed in Chapter 9 of this study: To what extent must certification requirements (which are largely synonymous with graduation from approved teacher training programs) be demonstrably related to effective performance on the job?

CBE and the NTE

The NTE is an instrument primarily designed to measure the academic achievement of college seniors who have completed four years of teacher education. In some states, its ready availability as an "objective" standardized instrument for considering the relative credentials of applicants for teaching positions, inclined state and local officials to utilize NTE scores on a cut-off basis, as a certification requirement or employment selection device. During the pre-Davis era, the courts which considered these practices uniformly prohibited them. The courts held that, in the absence of validation studies relating their content to classroom teaching duties in the particular locale, the tests could not be considered job-related, and therefore their use would be proscribed. These holdings were bolstered by testimony from representatives of the Educational Testing Service, the creator of the NTE, who stated that use of the KTE for credentialing or hiring purposes, without any evidence of attempts to validate it for such use, is inconsistent with the purpose for which the examination was designed.

In these rulings, the courts specifically applied the EEOC Guidelines validation criteria. For example, in Walston, the district court judge, although agreeing that predictive validation of the use of the NTE as a hiring standard had not been attempted, upheld the exam on general content validity

grounds. He reasoned that, since evidence at the trial indicated that subject area knowledge made up "25 to 30% of composite teaching behavior," the NTE commons examination, which is based on the general teacher education curriculum, was "rationally related" to the job.[111] The Court of Appeals for the Fourth Circuit, however, rejected this "common sense," "flexible" approach, stating that without a proper job analysis or validation study, sufficient content validation could not be established. The appeals court indicated that it was not clear that the subject matter tested on the national NTE exam sufficiently corresponded to the curriculm actually taught in the local classrooms. Thus, consistent with the premises of CBE, the court's decision implied that the proper assessment of ability to engage in teaching should be based on specific, demonstrable job-related competencies, rather than on a standardized examination related to abstract college courses.*

The first reported NTE case decided since Davis, dramatically differed from this consistent prior pattern of rigorous application of specific validation requirements. In United States v. South Carolina,[113] a three-judge court for the first

*"If these questions are a fair example of the remainder of the examination, any connection between the examination and effective teaching is purely coincidental. The NTE, as used by the Board, does not purport to measure the teacher's actual knowledge of the subject matter assigned to be taught or his performance in the classroom, but places primary emphasis on general education and professional education." 492 F. 2d at 926.

time upheld the use of the NTE as a statewide certification requirement. On its face, this direct reversal of prior trends would appear to indicate a substantial impact of the "flexibility" message emanating from the Davis decision. However, consistent with the discussion, at p.[26], supra, South Carolina might also be seen as the first "second generation" NTE case. That is, unlike the defendants in all the prior situations, the state officials here had conducted a validation study (in fact, since the study had been undertaken under contract with the Educational Testing Service, the expert testimony of ETS in this case was offered in support of the defendant's position, rather than in opposition, as in the prior cases).

In considering the import of this validation study, it must be understood that it did not purport to measure actual job-related abilities or competencies of would-be teachers. Instead, the prime assignment of the educators, assembled as the validation panel by ETS was to assess "the content validity of the NTE as compared to the curriculum in South Carolina institutions."[114] In other words, the ETS, validation study did not claim to be analyzing the extent to which the NTE was related to on-the-job performance of teachers in South Carolina schools, instead, it attempted to demonstrate that the questions on the national exam fairly represented the content of the curriculum being taught in South Carolina's teacher training institutions. Thus, in upholding the use of the NTE on the basis of this study, the court was accepting a "training course

validation" approach inconsistent with the specific job performance validation requirements of both the EEOC and the FEA guidelines in effect at the time.[115]

The South Carolina court gave little attention to the specific requirements in the applicable guidelines* and relied directly on Davis, noting that the Supreme Court there had upheld similar training course validation "wholly aside from its possible relationship to actual job performance as a police officer."[116] Despite the obvious similarities, however, the South Carolina court's ruling clearly represented a substantial extension of the Davis holding. The Supreme Court there upheld training course validation for a minimal proficiency verbal ability test, not for a detailed certification requirement like the NTE. Furthermore, Test 21 in Davis was held acceptable under specific civil service regulations in a context where Title VII did not directly apply; South Carolina upheld training course validation in a situation where Title VII fully applied.

*Somewhat inconsistently, the court stated that expert testimony at the trial, including that of Dr. Robert M. Guion, principal author of the APA standards, indicated that the ETS study design met all the requirements of the APA standards and the EEOC Guidelines, but then stated "To the extent the EEOC Guidelines conflict with well-grounded professional standards, they need not be controlling." 445 F. Supp. at 1113. These somewhat confusing statements presumably indicated that the procedures used in the validation study complied with APA and EEOC requirements, although these procedures were geared to training course content rather than on-the-job duties, would not be in compliance with EEOC standards.

Thus, in comparison with the previous line of NTE decisions which gave direct support to the competency assessment approach to certification, the South Carolina case appeared to endorse a more traditional course content assessment approach. In the long run, the South Carolina ruling, even if followed by other courts, could also be supportive to the CBE movement. Relevant in this regard is the fact that, at several points, the court emphasized the significance of the fact that no viable alternative to the use of the NTE had been suggested by the plaintiffs in the case:

> Here, plaintiffs have suggested only one alternative to the use of the NTE for certification purposes. Plaintiffs contend that mere graduation from an approved program would be sufficient and would have a lesser disparate impact on blacks. We cannot find that this alternative will achieve the state's purpose of certifying minimally competent persons equally well as the use of a content validated standardized test. The record amply demonstrates that there are variations in admission requirements, academic standards and grading practices at the various teacher training institutions within the state. The approval that the state gives to its teacher training programs is to general subject matter areas covered by the program, not to the actual course content of the program, and not to the means used within the program to measure whether individual students have actually mastered the course content to which they have been exposed. The standardized test scores do reflect individual achievement with respect to specific subject matter content which is directly relevant to (although not sufficient in itself to assure) competence to teach...[117]

Thus, the Court here was not approving the traditional "program approval" approach to teacher certification. On the contrary, it was only because of the very deficiencies of this approach that the standardized NTE was considered the preferred,

practical alternative.* This perspective leaves open the possibility that if a plausible alternative competency assessment model had been shown to exist, the court may have more strictly applied the letter of the Guidelines in requiring full job performance validation. As the court noted:

> The statistical studies in the record do not prove that high NTE scores would correlate with high scores on measures of teaching effectiveness. However, all the experts who testified in this case, including those offered by the plaintiffs, agreed that there is as yet no satisfactory measure of teaching effectiveness. For that reason, evidence of a lack of correlation is unpersuasive.[118]

In short, under the facts of the particular case, the court upheld the competency assessment approach that appeared to be the most rigorous available under the present state of the psychometric art. Seen in this light, the decision would not preclude judicial enforcement of more demanding job-related validation standards in this area, if it were plausibly shown that techniques for predictive validation or improved content validation of job-related teacher competencies were actually available.**

---

*South Carolina, in its consideration of this point, may also be viewed as a "diploma" case. Seen in this light, the Court can be said to have taken a strong position against certification based on invalidated diplomas issued by colleges of widely varying quality and to have held that states will be required to utilize more rigorous, standardized examinations.

**Note in this regard, the continuing insistence even in the Fourth Circuit (the Circuit covering South Carolina) of adherence to EEOC standards in the NTE context. Thus, in December 1977, after Davis, the Appeals Court rebuked the district court in Walston for not properly applying its previous order, and again reiterated that use of the NTE should be prohibited, even if it were not the sole selection device, without

This interpretation of the implications of South Carolina is not inconsistent with the Supreme Court's decision to affirm, by a slim, five justice majority (and without opinion), the lower court's ruling.* This summary affirmance may well reflect a view by the Supreme Court majority that, as in Davis, particular facts justified upholding defendants' practices,[119] but that a full analysis of the applicability of Title VII standards in this context should not be undertaken just because the Court did not want to undermine the fundamental, continuing applicability of Griggs and Albemarle.

The unwillingness of the Supreme Court to endorse generally the use of validation methods inconsistent with Griggs and the Uniform Guidelines is further indicated by the dissenting opinions of Justices White and Brennan in South Carolina. (Their dissenting comments constituted the only written statements to emanate from the Court in this case.) Justice White expressed in strong terms his belief that the lower court in South Carolina had misapplied the Davis holding on the training course validation issue:

> Washington v. Davis, in this respect, held only that the test, which sought to ascertain whether or not the applicant had the minimum commmunication skills necessary to

---

a demonstration of "proper validation studies and job analysis." Walston v. County School Board, 566 F. 2d 1201 (4th Cir. 1977).

*The Supreme Court's summary acceptance of the lower court's decision is technically an affirmance of the result, but not of any particular reason stated by the court below.

> understand the offerings in a police training course, could be used to measure eligibility to enter that program. The case did not hold that a training course, the completion of which is required for employment, be validated in terms of job-relatedness. Nor did it hold that a test that a job applicant must pass and that is designed to indicate his mastery of the materials or skills taught in the training course, can be validated without reference to the job.[120]*

Although Justice White was in the minority in South Carolina, his interpretation of the meaning of Davis on this point must be given great weight; he was the author of the majority opinion in Davis and wrote the specific sentences which were relied upon by the lower court in South Carolina.**

---

*In other words, there is an important distinction between an assessment instrument that is used to measure minimal ability to understand and successfully complete the curriculum in a training program (Test 21), and an assessment instrument that is used to measure ability to actually perform complex duties on the job (NTE). Extending the logic of Justice White's distinction, one might argue that the proper application of the Test 21 analogy to the teacher certification process would be to only allow use of a non-performance validated verbal ability test (or SAT aptitude test?) for selecting the freshman class at teacher training programs.

**It is also important to note in this regard that the trend to date, in the other lower federal courts which have considered training course validation issues under Title VII since Davis has been to continue to require a showing that the content of the training course itself is job-related. See, e.g., Blake v. City of Los Angeles, 595 F. 2d 1367, 1382 n. 17 (9th Cir. 1979); Dickerson v. U. S. Steel, 17 EPD 6694, 6727-8 (E.D. Pa. 1978); NAACP Ensley Branch v. Siebels, 13 EPD 6793, 6799 (N.D. Ala. 1977); Guardians Association of New York Police Department v. Civil Service Commission, 431 F. Supp. 526, 546-9 (S.D.N.Y. 1977), vacated and remanded, 562 F. 2d 38 (2d Cir. 1977); on remand 466 F. Supp. 1273 (S.D.N.Y. 1979); U.S. v. City of Chicago, 549 F. 2d 415, 431 (7th Cir. 1977). Contra, Richardson v. McFadden, 549 F. 2d 744 (4th Cir. 1976) (training course validation upheld in non-Title VII situation); U.S. v. Commonwealth of Virginia 545 F. Supp. 1077, 1100 (E.D. Va. 1978).

In McFadden, although the court upheld training course validation for a bar examination under constitutional standards, it specifically indicated that "if we were to determine Title VII standards were applicable, it would be necessary to reverse

Performance Evaluation

As discussed in Chapter 4, competency-based education emphasizes a variety of assessment techniques and, in particular, often substitutes performance evaluation for traditional paper-and-pencil testing procedures. The courts have not, to date, directly reviewed the legality of competency-based education models, but they have been called upon to review performance evaluation approaches in a wide variety of contexts,* including educational settings. For example, in Chance v. Board of Examiners,[121] after the traditional examination system, based largely on written tests, was invalidated, the court adopted an interim on-the-job performance evaluation system, which has now been in effect in the New York City school system for more than five years.

---

and declare the South Carolina Bar Examination constitutionally invalid." 540 F. 2d at 747.

The McFadden court's skepticism concerning the value of training school curriculum validation for a licensing examination is further indicated by the following interesting footnote:

> If it were carried to its logical extreme, seldom the path of the law, the Court's opinion on this point surprisingly might invalidate almost all state professional examinations. If the only demonstration of job-relatedness required is that it has a positive relationship to training course performance (e.g. law school), then why does not training school performance itself demonstrate that the applicant is fit to practice his profession? It is certainly clear that nothing correlates better with training school performance than training school performance itself. An applicant for the Bar who has graduated from an accredited law school may arguably be said to stand before the examiners armed with law school grades demonstrating that he possesses job-related skills. Why, then, any bar examination at all?

*Generally, in recent years, there has been a marked shift in the types of jobs at issue in Title VII cases, away from

As in Chance,* the courts generally appear to be receptive to the good faith implementation of new performance evaluation techniques, especially where a consensus of professional opinion appears to support their validity, or at least their value as positive experimental techniques. For example, although no court has yet specifically analyzed the "assessment center" approach to competency evaluation, and ruled on its validity,

---

lower-level blue-collar positions (as in Griggs) and toward higher-level managerial and professional jobs which rely less on traditional paper-and-pencil testing. See Stacy "Subjective Criteria in Employment Decisions Under Title VII," 10 Ga. L. Rev. 737 (1976); Note, "Title VII and Employment Discrimination in Upper Level Jobs," 73 Colum. L. Rev. 1614 (1973).

In a recent survey of 139 companies, 90% utilized formal performance evaluation programs for supervisors, middle managers and professional-technical personnel. Fifty-nine percent used such techniques for production workers, and 80 percent for office and sales personnel. Cited in Holley and Field, "Performance Appraisal and the Law," 26 Lab. L. J. 423 (1975).

*In Chance, Judge Mansfield originally expressed great skepticism as to whether any type of examination could be constructed to adequately assess the complex duties of the school supervisor, which depend "not so much on his knowledge of duties and educational content of courses given by his subordinates, as on such intangible factors as leadership skill, sensitivity to the feelings and attitudes of teachers, parents and children, and ability to articulate, to relate, to organize work...". 330 F. Supp. at 217. Notwithstanding these doubts, however, both parties in the case agreed that adequate performance assessment measures could be devised, and Judge Mansfield accepted their perspective on this point. See also Afro-American Patrolmen's League v. Duck. 366 F. Supp. 1095, 1103 (N.D Ohio 1975). Similarly, in Arnold v. Ballard, 390 F. Supp. 723, 731-2 (E.D. Ohio 1975), after initial invalidation by the court of traditional, written examinations for policemen, the city implemented, with the court's blessings, a sophisticated, professionally developed performance evaluation system based on a modification of the Landey-Heckman scales for rating the abilities of police officers.

that approach has been noted with favor in several cases.[122]

Where, however, performance evaluations were undertaken in a context which did not denote good faith, or were strongly criticized by the weight of professional opinion, the courts have not hesitated to invalidate the practices. The most prominent case in point in this regard was, of course, Albemarle,* where the Supreme Court refused to accept the results of "subjective, supervisorial rankings." In that situation, supervisors were asked to determine within each of a number of job groupings which employees were doing a better job than the person they rated against. The Court held that this procedure was in violation of Section 1607.(b)(3) and (4) of the EEOC Guidelines because:

> There is no way of knowing precisely what criteria of job performance the supervisors were considering, whether each of the supervisors was considering the same criteria, or whether, indeed, many of the supervisors actually applied a focused and stable body of criteria at any time.[124]

The other leading precedent most often cited in judicial discussions about performance rating techniques is Rowe v. General Motors Corporation.[125] There, the company's promotion standards were invalidated by the court under Title VII job-relatedness standards, because the system basically relied upon judgments by foremen who were given no written instructions, operated under vague standards, and the system as a whole afforded no procedural safeguards. In other cases, the Fifth

*See discussion at p.[11 ff] supra.

Circuit (which, having jurisdiction over the deep South, has apparently reviewed the largest number of these cases) criticized evaluation ratings, even when written instructions and criteria were provided, where the ratings were based on such general characteristics as "leadership, public acceptance, attitude toward people, appearance and grooming, personal conduct, outlook on life..." which the court considered to be "subject to partiality and to the whim of the evaluator."[126]

The courts have recognized, however, that in any system of performance evaluation, a certain degree of subjectivity is inherent. Thus, distinguishing between "vague subjective criteria" and "job-related subjective criteria," the district court in Nath v. General Electric Company[127] included among acceptable criteria such factors as "experience," "service," "results," " skills," "adaptability" and "versatility."

It may, of course, be difficult to understand why general items such as "adaptability" or "versatility" are considered acceptable, "job-related," subjective criteria, while items such as "leadership" and "verbal expression" are not. Although arguments concerning the comparative susceptibility of these items to concrete definition and application might be made, the real distinction in judicial attitudes and holdings appears to lie more in the degree of good faith, conformity to professional standards, and use of objective review procedures, than in the particular substantive criteria used.[128]

This conclusion may be illustrated by reference to the Supreme Court's reluctance to scrutinize clinical performance

evaluation criteria utilized by the medical school faculty in Board of Curators of the University of Missouri v. Horowitz.[129] In Horowitz, before accepting the recommendations of its own clinical faculty that a medical student should be dismissed from the school for substandard performance, the university had empanelled a special group of seven physicians experienced in the area, each of whom was asked to rate the plaintiff's performance in a variety of areas. Five of these seven doctors confirmed the negative ratings. Moreover, the plaintiff was afforded notice of her negative ratings upon several occasions, and was given ample opportunity to discuss and dispute the evaluation results. The Court was highly impressed with the fairness of the procedures accorded to the plaintiff, and even Justice Marshall, who dissented from portions of the majority opinion, stated that "it was difficult to imagine a better procedure" than the one used here.[130]

The contrasting results in the Supreme Court's decisions in Albemarle and Horowitz may be viewed as two poles of a continuum of judicial attitudes toward performance evaluation. Where the evaluators' motives and good faith are suspect (especially in a context where overt, intentional racial discrimination had existed in the recent past), their techniques are not in accord with established professional standards (or in violation of specific administrative guidelines), and concrete due process appeal procedures are lacking (all of which occurred in Albemarle), the Court will substantially scrutinize the evaluation. On the other hand, where the good faith of the

evaluators is not seriously in issue, practices appear to conform to professional standards and objective review procedures have been afforded the plaintiff (all of which occurred in Horowitz), the Court will not attempt to second guess the evaluators' substantive judgments.*

Generally, judges do not feel well qualified to assess the substantive criteria used in a performance evaluation process. Consequently, in those cases where they find themselves legally compelled to undertake such assessments, they are likely to rely on standards that are readily adapted to the procedural tools of the judicial...decision-making" process.** Specifically, they look to concrete statutory or regulatory provisions

---

*In general, absent a history of overt racial discrimination, courts tend to be deferential to judgments of school administrators, see, e.g., Ingraham v. Wright 430 U.S. 651 (1977), perhaps more so with college level administrators than with elementary or secondary school officials. See Goss v. Lopez, 419 U.S. 565 (1975) (high school suspension procedures subjected to judicial scrutiny), Larry P. v. Riles, p. 72 ff infra.

**Horowitz, supra, note 129, 435 U.S. at 90. See also James v. Stockham Valves & Fitting Co, 559 F. 2d 310 396 (5th Cir. 1977). Note in this regard the general unwillingness of courts to second guess evaluative judgments in the numerous recent cases involving denial of tenure decisions. See, e.g., McEnteggard v. Cataldo, 451 F. 2d 1190 (1st Cir. 1971) (tenure denial based on rater's judgment that a teacher was "difficult to get along with" upheld). See also, Holley and Field "The Law and Performance Evaluation in Education," 6 J. L. & ED. 427 (1977). Since most universities have adopted detailed formal procedures involving departmental, deanship, and presidential reviews before rendering final tenure decisions, courts generally are inclined to accept whatever judgment emerges from this objective process. Those few plaintiffs who prevail in these cases normally are awarded a right to a formal administrative hearing rather than reinstatement per se. See, e.g., Board of Regents v. Roth 408 U.S. 564 (1972), Perry v. Sinderman, 408 U.S. 593 (1972).

(e.g., the EEOC Guidelines) or to objective procedural requirements, such as due process guarantees. Thus, those competency-based education and perfomance assessment measures which are consistent with acceptable professional standards, are implemented in good faith, and provide reasonable procedural guarantees to those being assessed, are not likely to be subjected to further probing judicial scrutiny.

I.Q. Tests and Ability Tracking*

Despite the Supreme Court's indications in Horowitz, supra, that the courts would be reluctant to scrutinize academic judgments of school officials,[131] the federal courts in recent years have, in fact, actively intervened on a number of occasions to prohibit or regulate use of I.Q. tests for school placement purposes. The circumstances in which such intervention occurred generally were consistent with the criteria for judicial scrutiny outlined, supra, at p. [69]. For example, where ability groupings were instituted by formerly segregated school systems in the deep South as an apparent subterfuge to continue patterns of racial segregation within nominally integrated public schools, the Fifth Circuit established an absolute prohibition against ability testing and tracking until such time as meaningful unitary systems had been fully established.[132]

*For a full understanding of the legal issues discussed in this and the following two subsections, the reader is encouraged to refer to the discussion on the "middle-ground" equal protection test set forth as an appendix to this chapter.

In other cases, outside of the deep South, the courts have invalidated I.Q. testing and ability tracking where discriminatory impact on racial minorities was found, adequate procedural protections were not available, and professional justification for the practices was not clearly evident. The leading case in this regard is Larry P. v. Riles, [133] a challenge to the practice in the San Francisco school system (and ultimately throughout the State of California) of assigning students to classes for the mentally retarded, primarily on the basis of their scoring below 75 on standardized I.Q. tests. The court found in Larry P. that, although blacks constituted 28.5% of the student population in the San Francisco school district at the time the suit was commenced, they constituted 66% of those assigned to classes for the Educable Mentally Retarded (EMR).

When this case was first considered by the Court in 1972, it issued a preliminary injunction against the use of standardized I.Q. tests for classifying and placing students in EMR classes. This decision was based on constitutional equal protection considerations which utilized the precedents in Griggs and other employment discrimination cases in concluding that tests which are not validated for the purpose for which they are utilized are not substantially related to a valid governmental purpose.

By the time a full trial was held in the case and a decision on the merits issued in 1979, a number of new statutory provisions had been enacted which specifically required that tests

used to classify handicapped students be validated and free of cultural bias.* Therefore, the court sustained the plaintiffs' claims on three alternative grounds: (1) California's policies and practices were intentionally discriminatory (strict scrutiny equal protection); (2) were not substantially related to a legitimate state purpose ("middle test" equal protection); and (3) they caused racial segregation in special education classes without adequate justification (federal Title VI, Rehabilitation Act and Education of the Handicapped Acts; and the California constitution's equal protection clause).

In its decision, the court strongly grounded its analysis in the analogous standards and precedents developed in employment testing litigation. Significantly, it adopted the three-part analytical format from Griggs and other Title VII cases. That is, it first evaluated statistical evidence and concluded that plaintiffs established a prima facie case that a disproportionate number of minority children were being placed in EMR classes. Then it shifted the burden to the defendants to show that the placement actually were based on criteria that were validly related to the educational needs of the children. Finally, it looked at evidence to determine whether there were practical alternative selection devices, i.e., ones that could serve these legitimate educational purposes without causing segregation. Two of the psychometric standards the court borrowed

*See, e.g., 20 U.S.C. S1412(2) (D) (5); 45 C.F.R. S89.35(b); 35 C.F.R. 121a.532(a).

from employment testing law and applied in special education context were the federal guidelines that insisted on empirical proof of validation, rather than mere reliance on a test's general reputation,[134] and the special rules applicable to cases of differential validation.[135]

Translating the employment discrimination testing guidelines into the education context presented certain conceptual problems. The central question was what proof of validation would satisfy defendants' burden to rebut the prima facie case of discrimination. Analogizing from the concept of "job-relatedness," the defendants argued that they proved validation by showing correlation of I.Q. scores with student "performance" in school, as measured by the criteria of achievement test scores and grades. But the court rejected this whole approach,* stating that:

> [Th]e notion of predicting 'job performance' cannot be effectively translated into the educational context... If tests can predict that a person is going to be a poor employee, the employer can legitimately deny that person a job, but if tests suggest that a young child is probably going to be a poor student, the school cannot on that basis alone deny the child the opportunity to improve and develop the academic skills necessary to success in our society.[136]

In other words, correlations of test scores and school performance did not justify the particular uses to which the scores were being put, given the defendants' duty to provide children with appropriate education. The criterion for validation in

*Alternatively, it held that such correlation with achievement scores was not probative, and proof of correlations with grades of black students was unconvincing. Slip op. at 70-72.

this situation must be defined in terms of the particular psychological characteristics that demonstrate that a child cannot benefit from instruction in the regular school program. Validity, defined, had not been proven.[137]*

In deciding that placement procedures were not validated, the court distinguished Davis and South Carolina. As to Davis, it rejected defendants' contention that the Supreme Court meant to endorse generally the use of standardized intelligence tests, and held that the specific Davis ruling upholding performance in a training course as a criterion to validate selection by intelligence test scores, was simply inapplicable to the special education practices challenged in Larry P.

In South Carolina, not only had defendants actually produced reasonably adequate validation studies, but also the plaintiffs had failed to show that there was a feasible alternative means to serve the legitimate and important educational goal of selecting qualified teachers. Finally, in both Davis and South Carolina, the defendants had acted in good faith, without discriminatory intent, whereas in Larry P. there was a history of discriminatory intent.** Thus, Larry P. exemplifies the continued relevance of Title VII validation standards in

*The Court indicated (at n. 84) that these characteristics probably called more for a construct validation rather than a predictive or content validation approach.

**Although the equal protection intent finding was not technically relevant to the validation analysis, we have noted that a background of intentional discrimination may cause a court to apply technical psychometric guidelines more rigorously. See discussion at p. [69] ff, supra.

suits challenging educational policies outside the area of employment, although the court here emphasized that their actual application to educational testing practices requires an awareness of the unique factors at play in the particular context.

In other cases where circumstances, similar to those involved in Larry P., were at issue, testing and tracking practices were also proscribed. For example, in Diana v. State Board of Education,[138] individual I.Q. tests were required to be normed for use with Mexican-American students, and in Hobson v. Hansen[139] the court held that expert opinion indicating standardized aptitude tests are "worthless because evident cultural bias was "persuasive."[140] The court in Hobson emphasized the importance of norming ability tests to the local, largely minority student population, instead of to the national, predominantly white, middle class population, which was the typical norm for standardized I.Q. tests.

The holdings in these cases apparently influenced Congress and the Department of Health, Education and Welfare to take a strong stand against the use of unvalidated ability-testing mechanisms in the eligibility criteria for school districts which seek funding under the Emergency School Aid Act. Thus, school boards seeking funding to support desegregation programs under the Act are required to show that any ability grouping is "based upon nondiscriminatory, objective standards of measurement which are educationally relevant to the purposes of such grouping..."[141]

Because of the widespread use throughout the country of standardized I.Q. testing and tracking systems,[142] and because it is generally acknowledged that these tests are inherently culturally biased against minority students from low socio-economic backgrounds,[143] it is likely that ability testing and tracking would be invalidated on a wide ranging basis if the courts were to closely scrutinize these practices across the board. To date, the courts clearly have not done so. In this area, as in other areas of educational policy, the courts maintain a reluctance to interfere with the academic process, barring the presence of a constellation of compelling factors that call for judicial intervention.* Thus, for example, in Berkelman v. San Francisco Unified School District,[144] the same court which had upheld the preliminary ban on I.Q. testing in Larry P., refused to strike down a system for admissions to a special academic high school, which, being based on academic grades, adversely affected certain minority groups. The court's opinion was clearly influenced by the district's good faith efforts in instituting a special affirmative action program for minority admissions, and by the fact that the admissions system substantially furthered the important purpose of upgrading education.

*Note in this regard that the Court of Appeals in Hobson, although affirming Judge Wright's basic ruling, interpreted his decision as invalidating the previous tracking system, but as not precluding future testing and tracking practices, which presumably would be undertaken in a manner that would eliminate the most egregious violations of the old system. 408 F. 2d at 186-190.

Furthermore, it was pointed out that, unlike the situation in Larry P., no harm was inflicted on students who were not admitted to the special program.[145] Although the court's decision here preceded Davis, its analysis of the "facial validity" of the selection device at issue was analogous to the Supreme Court's consideration of Test 21. The court noted the obvious likelihood that those who had received high grades for academic achievement in junior high school would be most likely to perform well in the special academic high school.[146] The assumed correlation between academic performance in the two settings was apparent from a "common sense" perspective, although if one probed the validation issue to a deeper level, it might well be that the grading system at all levels of the school system was culturally biased against minorities, just as further probing in Davis might have indicated that the training program to which Test 21 was correlated itself was not job-related.[147]

An interesting comparison to the result in Berkelman is provided by the court's holding in Morgan v. Kerrigan,[148] where the admissions system for the elite Boston Latin Schools was substantially modified. The court indicated there that the Secondary School Admissions Test (SSAT), proposed by several parties as an objective, neutral admissions standard, was not validated, and undoubtedly would have disproportionate impact on minority students. Because of its findings of unconstitutional discrimination throughout the school system, the court

would not accept such a result. However, since the SSAT was apparently the "best available" alternative, the court permitted continued partial utilization of this mechanism (while also encouraging development of validated alternatives), provided that the overall admissions system was modified so that approximately 35% black or Hispanic students would be admitted.

The judicial approach toward the ability test-tracking issues in the reported cases has been based implicitly (and at times explicitly) on the application of the middle-ground "fair and substantial relationship" test discussed in the appendix, which allows for more extensive judicial scrutiny than has traditionally been applied to state actions challenged in constitutional equal protection grounds (but apparently less scrutiny than would apply under Title VII). For example, in its initial decision in Larry P., the court considered in great detail the question of what burden of proof should be placed upon the defendants, and concluded that, as in the employment discrimination cases, the clear pattern of discriminatory impact required more than the traditional rational relationship standards to be applied. (Under a rational relationship approach, the defendants' argument that I.Q. test screening was the "best available" methodology probably would have been considered acceptable.*) Its 1979 decision, although based on

*Similarly, the argument that providing education suited to the individual needs of handicapped children would require funding beyond the present ability of the school district

new statutory holdings which lessened the significance of this constitutional issue, specifically reiterated the continued validity of the middle-ground approach and held that Davis should not be interpreted as being inconsistent in this regard.[149]

Although denying relief to the plaintiffs, the court in Berkelman, supra, specifically adopted the "fair and substantial relationship" test. After citing a number of the leading "middle-ground" cases and articles, the court stated:

> Where a nonsuspect classification (past academic achievement) is alleged to operate to the detriment of a disadvantaged class or classes (black and Spanish-American students), neither "strict" nor "minimal" scrutiny provides useful guidance as a standard of review. The task is to examine the school district's assertion that the standard of past academic achievement substantially furthers the purpose of providing the best education possible for the public school students in the district.[150]

Similarly, in Board of Education, Cincinnati v. Department of H.E.W., the court summarized the cases in which tracking situations resulted in disproportionate impact on minority students and noted that the standard used was to require the defendants to show a "substanial congruence" between the testing process and the purposes of the assignment.[151]

---

probably also would have constituted a reasonable basis for school district exclusion practices under the rational relationship clause, but these arguments were explicitly rejected in Mills v. Board of Education, 348 F. Supp. 866 (D. D.C. 1972). See generally Note, Equal Protection and Intelligence," 26 Stan. L. Rev. 647, 659 (1974).

In short, the courts have repeatedly expressed skepticism about the validity of standardized aptitude tests for public school classification purposes,[152] but, in the absence of factors indicating strong discriminatory impact or actual discriminatory intent, apparent lack of conformity to professional standards,* and lack of objective review procedures, the courts are likely to restrain themselves from close scrutiny of particular practices.[153]

Student Competency Assessment

In the forefront of the education reform trends of the past decade has been the "accountability" movement. Fed by a variety of motives--holding schools and teachers responsible for low levels of student accomplishment, motivating student achievement, or assuring maximum return for burgeoning school tax payments--the common denominator of this trend has been an emphasis on measuring the end product of schooling, i.e., student performance. The initial legal attempts to promote accountability

---

*Note in this regard that since the enactment of the Education for the Handicapped Act (P.L. 94-142), and regulations promulgated thereunder, an extensive system of due process procedures for review of decisions concerning the placement of handicapped students has been put into effect in most states. The availability of these due process procedures may avoid the egregious results that have existed under past practices (retesting in Washington, D.C. indicated that two-thirds of those placed in special education classes were improperly assigned, and in Philadelphia, 25% of diagnoses were deemed erroneous, and another 43% questionable (Kirp, note 143, supra at 719).

emphasized novel "educational malpractice" theories. More recently, accountability concerns have centered on minimum competency testing (MCT) practices instituted by a large number of states.

Educational Malpractice

Educational malpractice claims are based on the proposition that many student learning deficiencies result from a school's failure to follow "generally accepted standards" of professional educational practice. Challenges to such professional negligence, in the form of suits for compensatory damages similar to malpractice claims historically brought under State Common Law against doctors, lawyers, or other professionals, were designed both to induce higher levels of professional performance and, at least in part, to provide funds to aid the victime to remedy his educational deficiencies, presumably by purchasing remedial services.

The highest appeals courts of California and New York have strongly rejected the idea of establishing such a right. These precedents are likely to exert considerable influence on any other state appellate courts that might be presented with this issue. The first major test case of the educational malpractice theory was Peter W. v. San Francisco Unified School District.[154] The plaintiff was a recent graduate of the school district who alleged that he had learned to read only at the fifth grade level after 12 years of schooling. He claimed that the school

was responsible for his reading deficiency because of its negligent failure to use reasonable professional care to diagnose his learning problems, to inform him and his parents of his relative achievement level, and to prescribe a proper course of instruction.

Accepting for purposes of argument the basic truth of these allegations, the court held that notwithstanding the school's moral and statutory obligations toward the plaintiff, a legal relationship supporting a claim for money damages did not exist. Lack of standards, it was said, necessitated this result:

> Unlike the activity of the highway or the marketplace, classroom methodology affords no readily acceptable standards of care, or cause, or injury. The science of pedagogy itself is fraught with different and conflicting theories of how or what a child should be taught...Substantial professional authority attests that the achievement of literacy in the schools, or its failure, are influenced by a host of factors...physical, neurological, emotional, cultural, environmental...[155]

In 1979, the New York Court of Appeals rejected two versions of educational malpractice claims. In Donohue v. Copiague Union Free School District,[156] a case with similar facts and claims to Peter W., the court unanimously reached the same legal conclusion as the California court. Later that year, in Hoffman v. Board of Education of the City of New York,[157] the judges voted 4-3 against accepting an asserted claim that was based on even stronger facts than either Peter W. or Donohue.*

*At the time of the trial, plaintiff was an adult who always had had at least average intelligence. Beginning at the kindergarten level, however, the public school professional staff

The New York court was relatively less concerned than the California court that adjudicating individual malpractice cases might be beyond judicial capabilities. It said:

> (T)he imagination need not be overly taxed to envision allegations of a legal duty of care flowing from educators, if viewed as professionals, to their students....Nor would creation of a standard with which to judge an educator's performance of that duty necessarily pose an insurmountable obstacle.[158]

Instead, the court's primary emphasis was on avoiding intervention into education policy-making.*

(relying on an erroneous I.Q. measurement) misclassified him as mentally retarded. They labeled him without any scientific basis as a mongoloid, and placed him in a class for the mentally retarded. This placement was based on scoring only one point below an absolute 75 cut-off demarcation. Although the initial school psychologist's report had specifically directed a re-evaluation within two years, in fact, no new test was administered for twelve years. (Such retesting is now required at least tri-ennually by federal regulations.) Throughout that time, Hoffman's teachers and psychologists attributed his low achievement to retardation. They never discovered that his learning problem actually was caused by speech and emotional problems. Looking back over this history, the intermediate appellate court had observed:

> So little had to be done to avoid the awesome and devastating effects of that failure on plaintiff's life, and that little was not done.

The Appellate Division accepted the plaintiff's argument that these facts constituted "affirmative negligence," which could be distinguished from a more generalized theory of educational malpractice, and it upheld a damage award of $500,000. 64 App. Div. 2d at 384.

*A number of other conceptual legal and educational policy objections have also been raised against educational tort theory. These include possible defenses of "contributory negligence," governmental immunity, difficulties in computing the dollar value of such damages, and possible negative impact on educational practices of a 'defensive' aura created by malpractice concerns. For a discussion of such issues, see Sugarman, "Accountability through the Courts," 82 Sch. Rev. 233 (1974).

> To entertain a cause for "educational malpractice" would require the courts not merely to make judgments as to the validity of broad educational policies--a course we have unalteringly eschewed in the past--but, more importantly, to sit in review of the day-to-day implementation of these policies. Recognition in the courts of this cause of action would constitute blatant interference with the responsibility for the administration of the public school system lodged by Constitution and statute in school administrative agencies.[158]

In short, no court has recognized a theory of tort liability for educational malpractice, and the unequivocal rejection by California and New York courts of the arguments in favor of such right make it unlikely that other state courts will entertain malpractice suits.*

---

*Two possible exceptions to this prediction are worth noting. First, some state courts may adopt the reasoning of the three members of the New York Court of Appeals who dissented from the dismissal of the Hoffman claim, even though they had agreed with the dismissal of the Donohue claim. They believed that a limited theory of negligent educational practice should be recognized and applied in egregious fact situations in which clear standards of professional practice have been breached.

Second, to the extent that courts are primarily concerned with questions of judicial manageability, as the California court seemed to be in Peter W., future judicial experience with minimum competency testing assessment measures may provide courts with the clear liability standards they now believe are lacking.

Minimum Competency Testing (MCT)

Minimum competency testing (MCT) refers to the use of standardized testing instruments to assess students' mastery of "basic skills." While standardized achievement tests have been used extensively for decades, the MCT programs currently being implemented or considered by virtually every state* represent a significant departure from past practices because: (1) standardized assessment measures are being implemented on a statewide basis; (2) criterion-referenced tests are used to measure whether students have attained behaviorally defined levels of academic achievement, or of "life skills;" and (3) test failure in some cases results in a major sanction, such as denial of a high school diploma regardless of a student's academic record.

There are, as well, unusual elements in the political genesis of MCT. The main impetus for this movement came from outside the ranks of professional educators, specifically from concerns of taxpayers and parents about the low level of academic skills of many high school graduates. Moreover, both the

*"As of March 15, 1978, 33 states had taken some type of action to mandate the setting of minimum competency standards for elementary and secondary students. All the remaining states either have legislation pending or legislative or state board studies underway." Pipho, "Minimum Competency Testing in 1978; A Look at State Standards," 59 Phi Delta Kappan 585 (1978); see also Chall, "Minimum Competency Testing in Reading: An Informal Survey of the States," 60 Phi Delta Kappan, 351 (1979).

MCT movement and reactions to it have led to novel configurations among traditional interest groups. Some conservatives have strongly favored MCT, at least in part because they believed that emphasis on basic skill tests will foster a return to traditional pedagogical methods. At the same time, some liberal educational reformers, including some minority group representatives, have strongly supported MCT, in the belief that dramatizing the degree of instructional failure of the schools will build support for more innovative educational approaches.

This contrast between liberal and conservative expectations reflects the fact that the MCT movement is a double-edged sword.* Widespread test failures may be read as indications either of student or of school system incompetence. Allocation of responsibility for failure can become a volatile issue, particularly when the system threatens to deny diplomas to those who fail. In response to concerns like these, several states have imposed on local school districts a duty to provide remedial instruction to students who fail their initial test administration.[160]

---

*It is also a double-edged sword from an educational perspective, on the one hand spurring ameliorative efforts to overcome diagnosed deficiences, but, on the other hand, sometimes lowering educational standards by promoting narrow "teaching to the test" techniques. See Madaus, "Testing and Funding: Measurement and Policy Issues" in W. Schrader, Ed., Measurements and Educational Policy (1979) at 56.

The widespread adoption of the MCT programs has spurred extensive commentary[161] in the legal literature, but so far only one major court decision, Debra P. v. Turlington.[162]

Debra P. was a class action brought on behalf of all present and future twelth grade students who failed or would thereafter fail the Florida functional literacy examination, a test of basic communication and computational skills. Although the students were given subsequent opportunities to retake the examinations (and in the meantime were provided special remedial programs), they were threatened with eventual denial of high school diplomas if they did not ultimately achieve passing scores. Plaintiffs also raised additional claims on behalf of two subclasses of black students who disproportionately failed the examinations. The adverse impact on black students was dramatic. For example, among students who had taken the test three times, the black students failed at a rate 10 times greater than did white students.[163]

After an extensive trial, the federal district court ruled that the State of Florida could continue to administer the tests and could assign students to remedial classes on the basis of the test results. But it enjoined the school authorities for four years from withholding a high school diploma from any student because of failure on these tests. There were three critical aspects of this decision.

(1) In regard to the subclass of black students, the court ruled that the diploma sanction amounted to an unlawful perpetuation of past discriminatory effects because black students

presently in high school had attended inferior, de jure segregated schools during their elementary school years.* (2) The court found that immediate imposition of the diploma sanction was arbitrary and a denial of constitutional rights to due process with respect to all high school students, black and white, because the sudden introduction of wholly new graduation requirements denied them of fair prior notice of the standards for obtaining a high school diploma and of a reasonable opportunity to prepare themselves for the new requirements.** (3) The Court upheld the content and construct validity of the test, after analyzing the content of the test under the "middle-ground" equal protection approach.

The court's discussion of the validation issues is particularly significant for purposes of this study. Citing a number of middle-ground equal protection precedents, the judge considered whether the content of the tests was substantially related to the state's legitimate educational objectives. He specifically noted the relevance of Griggs and Title VII employment discrimination cases in this context and utilized the

*The court rejected, however, the plaintiff's additional claim that the assignment of students to remedial classes based on their scores on the competency tests was unconstitutional resegregation. It also rejected the argument that exemption of private schools from the Minimum Competency Program and diploma requirements was racially discriminatory.

**In this part of its decision, the court distinguished the U.S. Supreme Court's decision in Horowitz, note 129 supra, indicating that courts will become involved in issues of academic judgment where broad legislative policies rather than individual student records are at issue, and where the plaintiffs are high school students attending under compulsory education laws, rather than graduate students. Slip op. at 40.

psychometric concepts established in these cases, but he did so in a "flexible" manner. Specifically, the court accepted as a given the definition of functional literacy established by the legislature without considering whether steps analogous to conducting an adequate "job analysis" under the EEOC Guidelines had been undertaken to establish the validity of these criteria. Accepting the criterion objectives as defined by the state, both content and validity was upheld for these tests.* Thus, although the court substantially considered the applicable psychometric concepts, it held that:

> ...the "state of the art" is not to be equated with the consitutional standards for Fourteenth Amendment due process and equal protection review.[164]

As a result of the detailed analysis of Florida's MCT program by the court in Debra P., minimum competency testing programs in other states should expect similar substantive judicial scrutiny of these programs if analogous legal challenges are asserted. The court's decision in the Florida case will provide

---

*In doing so, the court also rejected plaintiff's claims that the cut-off scores for determining the pass level on these examinations was arbitrarily established. This claim is related to a fundamental criticism that the concept of "minimal competency" cannot be fairly determined in an objective manner because any absolute definition of a minimum quantum of skills needed to function independently as an adult is per se arbitrary. See, e.g., Glass, "Standards and Criteria," 15 J. Ed. Measurement 237 (1978). Haney and Madaus, note 163 supra at 468, concluded:

> At present there simply is no scientific foundation for deciding what minimum points should be; the decisions involved in setting them are political rather than scientific.

strong precedent for invaliding diploma sanctions having adverse impact on minority students in states having a history of de jure school segregation.* More generally, the court's willingness to apply Title VII validation concepts (albeit in a "flexible" manner) to the content of the tests as they affect all students means that the validity of other MCT programs subject to future court challenges will need to be demonstrated, and the extent to which these programs generally are consistent with established psychometric practices will undoubtedly be a major factor in influencing the ultimate outcome of such cases.**

Graduate School Admissions Tests

Traditionally, decisions as to which of the thousands of applicants for the limited number of places available in

*In states not having a history of past de jure segregation, but whose MCT programs have a strong adverse impact on minorities, the claim may also be made that under the general anti-discrimination mandates of Title VI of the 1964 Civil Rights Act, defendants should be required to "validate" substantially any school policies that disproportionately affect minority group students in the same way that Title VII standards are invoked by discriminatory impact, whether or not accompanied by discriminatory intent. The Supreme Court's decision in Lau v. Nichols, 414 U.S. 563 (1974), would appear to lend force to this "impact" interpretation of Title VI. However, in Bakke, four Supreme Court Justices appeared to question the continued soundness of Lau following Davis, and suggested that Title VI, like the Constitution, may only apply in situations of intentional discrimination.

**Debra P. constitutes the first major testing case to apply middle-ground equal protection analysis to claims asserted by non-minority students. This aspect of the decision may also have important precedential implications in other educational testing contexts. See Appendix.

graduate and professional schools would be accepted for admission* largely were based on undergradute grade point averages and standardized admissions tests such as the Law School Aptitude test (LSAT) or the Medical College Admissions Test (MCAT). It has been generally acknowledged that in the absence of special affirmative action admissions procedures, these standardized admissions criteria would have a substantial detrimental impact on minority applicants.** In light of these realities, many graduate schools have established "race conscious" special admissions procedures to ensure an adequate representation of minority students in their student bodies.

Legal challenges to these affirmative action admissions procedures have generated enormous public controversy and commentary both in the popular press and the scholarly literature. The United States Supreme Court managed to avoid squarely facing these highly charged issues in 1974 by declaring the controversy concerning the law school admissions procedure at the University

---

*In University of the Regents of California v. Bakke, 438 U.S. 265 (1978), 3,737 students applied in 1974 for the 100 available places; in Alevy v. Downstate Medical Center, 39 N.Y. 2d 326 (1976), 6,300 candidates applied for 216 positions to the medical school program; in Defunis v. Odegaard 416 U.S. 312 (1974), 1601 applications were received for 150 available places.

**One extensive study found, for example, that without special affirmative programs, utilization of the "predicted first year grade point average" based on undergradute grades and Law School Admissions Tests would reduce black law school enrollment by 82%. Sindler, "Bakke, Defunis and Minority Admissions" 141 (1978). The lower Court decision in Bakke indicated that almost all minority candidates were rejected under traditional admissions procedures. 132 Cal. Rep. 680,712 (1976) (Tobriner, J., dissenting).

of Washington moot in light of the fact that the plaintiff, who was temporarily admitted pending his appeal, was about to graduate.[165] In 1978, however, the Court was called upon to squarely confront these issues in a case brought to challenge the medical school admissions system at the University of California, Davis.[166]

Under the Davis system, members of specified minority groups, who claimed to come from educationally or economically disadvantaged backgrounds, were permitted to have their applications reviewed by a special admissions committee. This special committee would recommend candidates for 16 of the 100 places available in the entering class. The minority candidates recommended by this committee generally had substantially lower undergraduate grade point averages and test scores than those admitted under the general admissions process. Allan Bakke, a white applicant who was denied admission, challenged the legality of this process, which he claimed precluded him from fairly competing for 16% of the available places in the entering class.*

The Supreme Court's treatment of this issue culminated in a lengthy, complex decision, containing six separate concurring and dissenting opinions. Basically, four members of the Court

*Interestingly, Bakke had both a higher grade point average and higher MCAT scores than the average of those students admitted to the 84 regular places in the class. Apparently he was denied admission through the regular admission process because of his comparatively low rating on the interview aspects of the admission process. Bakke, note 74 supra at 277, note 7.

stated that the preferential admissions system should be held to meet constitutional requirements, four members were of the opinion that Title VI of the 1964 Civil Rights Act precluded any type of race conscious admissions system, and would not reach the constitutional question, while the ninth member of the Court, Justice Powell, who thus became the "swing vote," held that the specific "quota" approach used at Davis was unconstitutional, but that other "race conscious" admissions systems might be constitutional.

The Court's complicated, compromise treatment of these issues, which probably would allow a university admissions committee to implement an affirmative action program assuring precisely the same minority representation, as did the one at Davis, but in a manner less likely to grate against majoritarian sentiments,* may be effective as a political solution to the immediate controversy.[167] However, the Court's hedged response leaves open both the basic constitutional issues raised by the case, as well as specific questions concerning the validity of the use of standardized admissions criteria for entrance to graduate school programs. Indeed, one of the most salient features of the lengthy Bakke decision is that in 156

*Note, "The Supreme Court," 1977 Term 92 Harv. L. Rev. 57, 146 (1978). Under Justice Powell's decision, a university seeking to "diversify" its student body could give extra consideration, on an individual basis, to ethnic background, and in assessing the extent to which such consideration would be appropriate, the admissions committee apparently could give "some attention to (total) numbers" 438 U.S. at 323.

pages of text no consideration whatsoever was given to the obvious, critical issue as to whether the standardized admissions procedures used at Davis were in any sense validated or reasonably related to the purposes for which they were being used. The reason for this anomaly was that the record in the case, as it came up from the lower court, did not present any specific facts on this issue. Plaintiff Bakke, of course, had no reason to challenge the validity of the standardized test scores, on which he performed relatively well. Similarly, the university officials, while defending the legitimacy of their special admissions program, also had no interest in raising questions concerning the validity of their general admissions procedures and their reliance on standardized testing techniques.*

---

*The University specificially justified its policies on four grounds: (1) increasing the traditionally low minority representation in medical schools and the medical profession; (2) countering the effects of societal discrimination; (3) increasing the number of physicians who will practice in currently underserved communities; and (4) obtaining the educational benefits of an ethnically diverse student body. 438 U.S. at 306. Justice Powell upheld the constitutionality of "race conscious" admissions practices only on the fourth count, student body diversity. In an interesting footnote, he observed that test validation problems might have been offered as a fifth justification:

> Racial classifications in admissions conceivably could serve a fifth purpose, one which petitioner does not articulate: fair appraisal of each individual's academic promise in the light of some cultural bias in grading or testing procedures. To the extent that race and ethnic background were considered only to the extent of curing established inaccuracies in predicting academic performance, it might be argued that there is no "preference" at all. Nothing in the record, however, suggests either that any of the quantitative factors considered by the Medical School were

If the Court had been in a position to directly consider the test validation issues under Title VII-type "job-relatedness" criteria, there is serious question as to whether the MCAT would have passed muster. A detailed brief setting forth the literature on validation of the MCAT was, in fact, submitted to the court on behalf of the Black Law Students Association at Berkeley as amicus curiae. The professional literature reviewed in this brief indicated that the MCAT, which was normed to the middle class medical school population in 1951, at a time when virtually no minority students were in attendance, is culturally biased against blacks and other minority groups. Although there was some evidence indicating a correlation between MCAT scores and course grades in the first two years of medical school, the evidence strongly indicated that the test was not validated either to grades in clinical courses during the latter

---

culturally biased or that petitioner's special admissions program was formulated to correct for any such biases. Furthermore, if race or ethnic background were used solely to arrive at an unbiased prediction of academic success, the reservation of fixed number of seats would be inexplicable. Id., at 306, n. 43

Justice Powell also noted in passing the relevance of Griggs and the Title VII cases to this test validation issue. (Id. at 307, n. 44), as did Justice Brennan joined by Justices White, Marshall and Blackmun in their decision (concurring in part and dissenting in part) Id. at 352 where, citing Griggs, they state:

> Although this Court has not yet considered the question, presumably, by analogy to our decisions construing Title VII, a medical school would not be in violation of Title VI under Lau because of the serious underrepresentation of racial minorities in its student body as long as it could demonstrate that its entrance requirements correlated sufficiently with the performance of minority students in medical school and the medical profession.

years of medical school, or to actual performance on-the-job after graduation. These conclusions have been corroborated both by critical commentators[168] and by evidence and findings of other courts.[169] In addition, the focus on graduate school admissions procedures in recent cases has also indicated a substantial degree of subjectivity and arbitrariness in the interviewing and decision-making process, factors which have generated strong judicial disapproval.*

Consideration of the validity of standardized admissions tests takes on special significance in the context of the highly competitive graduate school admissions process. Typically, from the large applicant pool, a small percentage of outstanding candidates at the top of the stack are immediately admitted, a small number at the bottom immediately rejected, and the large remaining group, all of whom are deemed qualified to successfully complete the course of study at the school, are left for selection decision.[170] Utilization of cut-off scores (whether or not on a "racially conscious" basis) to select an entering class means than critical career-determining decisions are made

*For example, in DeFunis v. Odegaard, 507 p. 2d 1169, 1193 (S. Ct. Wash. 1973), it appeared that two student members of the admissions committee who had full voting authority were selected because they had volunteered through the student bar association; no criteria concerning their qualifications for this job had ever been established. Each of these students was given approximately 70 files upon which to make initial recommendations for admission or rejection, with instructions that only ten were to be passed on for further consideration by the full committee. In other words, most of the applications were summarily rejected by individual committee members, including student members, with no further scrutiny by the full committee.

on the basis of narrow differentials which have not been validated for such a purpose. Although, as indicated above, some evidence indicates a correlation between standardized scores and medical school grades on course work in the first two years, the situation is substantially different from the "training course validation" upheld by the Supreme Court in Washington v. Davis. In Davis it was assumed that those who did not pass Test 21 were not capable of completing the training course curriculum. In the graduate school context, no such assumption can be made for most of the rejected applicants.

To date, the most detailed judicial consideration of the validity of standardized testing instruments for graduate school admission purposes was contained in the dissenting opinion of Justic Douglas in the DeFunis case. Justice Douglas dissented from the majority's decision to declare the controversy moot and went on to discuss in detail his perspective on the issues. In doing so, he questioned the usefulness of multiple-choice type examinations for evaluation of the creativity and intelligence of potential law school students, the cultural bias of the exams, as well as what he considered an over-reliance upon relatively small differentials in scores on these examinations.[171] He concluded that:

> [T]here is no clear evidence that the LSAT and GPA provide particularly good evaluators of the intrinsic or enriched ability of an individual to perform as a law student or lawyer in a functioning society undergoing change.[172]

However, according to Justice Douglas, the solution to the cultural bias and disparate impact of standardized graduate school admissions tests was not to adopt a numerical quota admissions scheme, but rather to substantially reform the admissions procedures so that they operate in a racially neutral manner which is fair to all:

> The key to the problem is consideration of such applications in a racially neutral way. Abolition of the LSAT would be a start. The invention of substitute tests might be made to get a measure of the applicant's cultural background, perception, ability to analyze, and his or her relation to groups. They are highly subjective, but unlike the LSAT, they are not concealed, but in the open. A law school is not bound by any legal principle to admit students by mechanical criteria which are insensitive to the potential of such an applicant which may be realized in a more hospitable environment. It will be necessary under such an approach to put more effort into assessing each individual than is required when LSAT scores and undergraduate grades dominate the selection process.[173]

Justice Douglas went on to propose for consideration broadened interviewing processes, both with the applicant and others who are in a position to assess his abilities, as well as the operation of summer school programs in which student performance could be directly measured. Justice Douglas' suggestions have apparently been adopted, at least in part, by one major university which presently conducts detailed assessment interviews for all those in the top ten percent of the applicant pool.[174] Indeed, presumably because of the critical focus on medical school admissions stemming from Bakke and other recent litigation, the Association of American Medical Colleges is in the process of substantially reforming medical school admissions procedures so that in addition to traditional subject area

knowledge, more general ability in written communications and problem-solving skills, as well as personal qualities deemed necessary for the successful practice of medicine, will be part of the assessment battery.*

In short, then, substantial questions concerning the validity of the use of standardized graduate school admissions tests are raised by application of Title VII-type standards in this area. Even though the courts have moved with extreme caution on this issue, the judicial "dicta" which have emerged, combined with the general awareness of validation requirements stemming from Griggs and its progeny, have already had a significant effect on reform trends in these areas. Because of the critical importance of these issues, especially where they have disparate impact on minority populations, more direct judicial consideration of the validity of standardized admission tests is likely to occur in the future.[175]

---

*An amicus brief submitted by the Association of American Medical Colleges reports that substantial general changes in medical school admission procedures may be imminent. 'Recognizing the current need for better assessment of academic and personal qualifications, the Association will begin use in 1977 of a revised set of admissions tests as part of a new admissions assessment program. Such tests will not only measure achievement in particular areas of knowledge pertinent to medical study, but will also demonstrate abilities and interpretation of written communications and in problem-solving skills. The development of instruments for evaluating personal qualities deemed necessary for the successful practice of medicine is underway. Seven broad areas have been identified for study: compassion, interprofessional relations, coping capability, sensitivity and interpersonal relations, decision-making capacity, staying power, and realistic self-appraisal.' pp. 9-10. Bakke v. Regents of University of California, 132 Cal. Rep. 680, 714-15, n. 15 (1976) (Tobriner, J., dissenting.)

III. CONCLUSION

Our detailed survey of the state of the law on competency assessment issues has yielded three fundamental conclusions.

1. The courts have become knowledgeable about sophisticated validation concepts and have repeatedly applied psychometric analyses to testing practices of employers and educators. In contrast to the traditional pattern of deference to the expertise of administrators and governmental officials, courts today manifest an ability and a willingness, under certain circumstances, to scrutinize testing practices.

2. Despite their knowledge and experience in this area, the courts nevertheless tend to act with restraint on these issues. Especially since the Supreme Court's 1976 decision in Washington v. Davis, the courts have been "flexible" in their application of the rigorous validation requirements set forth in the administrative regulations. This restraint and flexibility is even more pronounced in the judicial approach to educational assessment issues brought under the equal protection clause of the Fourteenth Amendment, where specific statutory mandates such as Title VII do not directly apply. In our review of the diploma, licensing and certification, performance evaluation and I.Q. testing cases, we detected a prevalent pattern, indicating that the courts generally would scrutinize testing practices only when one or more of the following factors was present: (1) lack of good faith of the defendants, especially in regard to possible

discriminatory motives; (2) substantial lack of adherence to applicable professional practices; or (3) denial of minimal procedural due process protections.

3. There is a significant interrelationship between the state of the practice in competency assessment techniques (at least as the state of the practice is made known to the courts), and judicial activism in this area. When they believe it appropriate to scrutinize the validity of challenged tests, courts will require substantial compliance with relevant validation standards, but as was dramatically shown in the South Carolina NTE case, they will not impose "ideal" or technical validation requirements which appear to be beyond the level of reasonable, current professional practice. Conversely, as illustrated by the Bridgeport police cases, court intervention on testing issues may have motivated a higher level of performance by employers and educators and may have stimulated progress in the development and application of fairer competency assessment measures.

The overall implications of these conclusions would appear to be that educators can assume that to some degree the courts are "looking over their shoulders" and they may expect to be held accountable for a reasonable measure of conformity to accepted professional practices in their testing procedures. However, professionals need not generally anticipate wide-ranging judicial overview of testing practices or radical decrees mandating implementation of theoretical or experimental validation standards.

Because of the complex, developmental nature of judicial activity in this area, we would recommend three specific areas of continuing legal research. First is the obvious need to continue to closely track judicial developments in this area: in Title VII cases; in each of the areas of application to education that were identified in part II of this chapter, and in other, newer areas of educational testing not covered in this study (largely because a body of legal doctrine has not yet emerged) such as special education, education of the gifted and talented, bilingual education and funding allocations.[176] From a technical legal point of view, it will also be important to closely follow the extent to which the "middle-ground" equal protection test may be generally expanded in major decisions of the Supreme Court, or in particular education decisions of the lower federal courts, including possible extension of middle-ground equal protection analysis, as in Debra P., to situations involving nonminority plaintiffs. The results of such research should, we believe, be disseminated to the educational community on a continuing basis so that major new decisions are immediately understood in proper perspective, and their implications are not distorted by impressions fostered in the popular press.

The second recommended area for further legal research concerns the decisions of the various state courts that bear on the testing issue we have considered. Because of the primacy of federal statutes (especially Title VII) and constitutional precepts, as well as the obvious difficulty of considering the

differing judicial systems of the 50 states in an initial legal overview, the present memorandum has exclusively considered the decisions of the federal courts. Since, despite recent federal initiatives, education remains fundamentally a state and local responsibility under the American federal system, it is obviously important to survey and analyze judicial developments in at least a representative number of state systems. At this point, we are not certain what such a survey would reveal. For example, the courts in New York State appear to be maintaining a traditional, highly deferential attitude toward the decisions of educational administrators on testing issues.[177] By way of contrast, the Supreme Court of New Jersey[178] has taken a more activist posture on funding and related educational competency assessment issues.

Finally, we believe that detailed research, on a case study basis, should be undertaken of the remedial phases of major educational testing cases. Almost all discussions of judicial involvement in educational policy matters, including the present chapter, are largely based on the pronouncements contained in the official decisions and orders issued by the courts. But developments in the remedial stage, where actual implementation of the decree is undertaken, are often of critical importance.* Many courts, in striking down defendant

*Little research is normally undertaken of these remedial processes because specific information can be gathered only by direct examination of voluminous court files and through interviews with those involved in the case, whereas judicial decisions are readily available in any law library. Although resort

testing practices, have mandated that future tests must be specifically validated in accordance with the EEOC guidelines.[179] In implementing such requirements, however, substantial practical problems often complicate the goal of achieving full scientific validation.

For example, in Chance v. Board of Examiners,* the ideal goal of replacing a nonjob-related examination system with one that was scientifically validated was compromised against political and practical considerations. The Chance court issued an injunction prohibiting the further use of existing Board of Examiners tests to license school supervisors and administrators in New York City. Under the pressure of this injunction, the parties to the case and other interested groups created an innovative new interim examination system based on initial screening by local community school boards. The persons hired were--after a trial period--evaluated under testing procedures carried out by Board of Examiners staff persons in consultation with community representatives. The interim system was never scientifically validated, but it resulted in the hiring of much larger numbers of minority educators than under

---

to individual case files may be overly time-consuming for the purposes of ordinary legal research, in the context of a social science analysis of implementation of testing reforms, these court files probably provide a more accessible, comprehensive source of rich primary data than almost anything else that is available.

*See discussion, pp. [65] supra.

the previous system and there was general agreement among most of the city's education constituencies that the new selection procedure worked reasonably well. This interim procedure was supposed to have been replaced by a more scientifically developed examination system involving entry level tests based on systematic job analyses. However, disputes arose about the adequacy of the prototype job analyses that were prepared by a consulting firm hired for that purpose by the Board of Education. Consequently, the essential elements of the interim system have been in use for six years as a de facto permanent system.[180]

As this example suggests, the process through which abstract judicial mandates have been transformed into actual institutional reforms constitute unique "laboratories" for studying public policy implementation problems. Certain reform approaches which are advocated in the literature may be found wanting, for various reasons, when actual attempts are made to put them into practice. On the other hand, the pressures of judicial oversight may, as in Chance, motivate the parties to devise and put into practice new techniques that previously had not been considered feasible. In short, we believe that detailed analysis of the remedial process of major court cases will provide valuable empirical data for further consideration of the critical issues of the apparent gap between ideal and actual practice, and the barriers to improved practice, revealed by the state of

the practice chapters of this study.*

*Because of the substantial gaps between actual practice and existing regulatory standards, we are reluctant to recommend any further statutory or regulatory reforms at this time. Further research into the remedial stages of past court cases, together with additional empirical research on the state of the practice as revealed in this study, may provide a substantial data base which could be presented to Congress and the regulatory agencies with appropriate recommendations for legislative action at some future date.

APPENDIX:

The Emerging "Middle-Ground" Equal Protection Test and Its Application To Education Law Litigation

The Emerging "Middle-Ground" Equal Protection Test and Its Application To Education Law Litigation

Judicial review of diploma requirements and educational hiring standards have largely been based on direct applicability of Title VII job-relatedness standards. As previously indicated,* these precedents may also be indirectly applied to situations involving standardized testing practices in a variety of educational contexts. However, no statutory mandate generally requiring validation of educational testing practices has been held to apply outside the employment discrimination context.** Accordingly, judicial consideration of the validity of I.Q. tests, minimum competency testing, and graduate school admissions exams have largely been based on general constitutional principles of equal protection of the laws, rather than on specific statutory or regulatory standards. In order to understand how courts have assessed claims of unfair testing treatment under the equal protection clause, and how they may assess such claims in the future, it is necessary to briefly consider emerging new trends in equal protection doctrine.

---

*See discussion, [p. 36 ff.] supra.

**See, however, indications in Regents of the University of California v. Bakke, 438 U.S. 265 (1978) discussed at pp. supra concerning the possible applicability of Title VI of the 1964 Civil Rights Act in this regard. Note also the specific requirements for tests resulting in classifications of handicapped students discussed at note 75 supra.

The development of constitutional equal protection doctrine from the New Deal era through the late 1960s resulted in two general approaches to laws and regulations whose impact falls unequally on differing groups of citizens: (1) in some cases such state actions would be analyzed to determine whether they had any "rational relationship" to a valid state purpose; while (2) other state actions would require a showing of a "compeling state interest" justifying the contested practice. Application of the "rational relationship" test imposes a minimal burden of justification upon public agency defendants in a law suit, since the standard would be met if "any state of facts reasonably may be conceived to justify it,"* even if other means with less burdensome consequences for affected groups could have been devised. If, on the other hand, the "compeling state interest" test is applied, the courts will closely scrutinize the practice at issue, placing on the state the heavy burden of establishing that no other available legislative or administrative methods could have achieve the desired result.** In fact, only in two cases, both dealing with the confinement of Japanese-Americans under the emergency conditions of World War II, has the Supreme Court upheld state actions which were subjected to this

---

*McGowan v. Maryland, 366 U.S. 420, 426 (1961).

**For a general overview of these points, see Note, "Developments in the Law: Equal Protection," 82 Harv. L. Rev. 1065 (1969); Michelman, "Foreword: On Protecting the Poor through the Fourteenth Amendment," 83 Harv. L. Rev. 7 (1969).

rigorous "compeling state interest" test.*

Because of the dramatic difference in outcome resulting from application of these differing equal protection standards, the key issue in many constitutional litigations is which standard should be applied.** Historically, the Supreme Court has applied the "compeling state interest" test only to state actions involving "suspect" classifications (usually classifications based on a race)*** or to cases involving certain defined "fundamental interests," such as voting rights,**** rights to a fair trial,***** and other specific constitutionally-identified areas such as the right to interstate travel.****** The Supreme Court recently held that education is not a "fundamental interest" under the federal constitution because, since education is primarily a state obligation, it is not one of the federal rights specifically protected therein.*******

---

*Korematsu v. United States, 323 U.S. 214 (1944); Hirabayashi v. United States, 320 U.S. 81 (1943). Justice Powell's decision in Bakke, supra also upheld race-conscious medical school admissions practices under the compeling state interest test.

**See, e.g., San Antonio Independent School District v. Rodriguez, 411 U.S. 1, 16-17 (1973).

***See, e.g., Loving v. Virginia, 388 U.S. 1, 8-9 (1967).

****See, e.g., Harper v. Virginia State Board of Elections, 383 U.S. 663 (1966); Dunn v. Blumstein, 405 U.S. 330 (1972).

*****See, e.g., Griffin v. Illinois, 351 U.S. 12 (1956).

******Shapiro v. Thompson, 394 U.S. 618 (1969).

*******San Antonio Independent School District v. Rodriguez, supra. Education has, however, been classified as fundamental interest under some state constitutions. See, e.g., Serrano v. Priest, 96 Cal. Rptr. 601, 487 P. 2d 1241 (Cal. 1974).

It would appear, therefore, that cases involving challenges to educational practices, which do not involve clear findings of discriminatory intent against minority groups* would call for application of the lesser "rational relationship" standard, which would involve minimal scrutiny by the courts and almost automatic approval of defendants' practices. However, as indicated in Section II of the main text, courts have in fact subjected educational testing practices to more than minimal rational relationship review, without invoking the rigorous, compeling state interest standard. These holdings have explicitly or implicitly been based upon a "middle-ground" standard of equal protection review, which has been developed in recent years, primarily in sex discrimination cases. This approach calls for greater scrutiny than would normally be applied under the minimal rational relationship test, without requiring a showing of compeling state interest or the lack of any conceivable alternative method for accomplishing a legitimate goal.

The Supreme Court itself has explicitly endorsed the "middle-ground" approach in recent years only in cases involving allegations of sex discrimination.** One gets the distinct

---

*See Washington v. Davis, supra.

**Craig v. Boren 429 U.S. 190 (1976), see also Reed v. Reed 404 U.S. 71 (1971), Frontiero v. Richardson 411 U.S. 677 (1973), Weinberger v. Wiesenfeld 420 U.S. 636 (1975). Some decisions in cases involving illegitimacy, age discrimination and alienage also appear to implicitly adopt middle-ground analyses, although the Court has not explicitly articulated the concept in these areas. See e.g., Trimble v. Gorden 420 U.S. 762 (1977). See generally L. Tribe, American Constitutional Law 16.31 (1978).

brought under general constitutional equal protection standards. Under traditional equal protection doctrine, "strict scrutiny" was applied only in cases involving apparent discriminatory intent and the courts were reluctant to extend that rigorous degree of scrutiny in the absence of clear indications of such motives. On the other hand, especially in light of the Title VII discriminatory impact standard contemporaneously being applied to actions of private employers, the courts were reluctant to use a minimal rational relationship approach. Thus, in one of the first cases to directly confront this problem, the district court in Chance took a middle position, stating that in the employment discrimination context, defendants should be required to make a "strong showing" to justify their practices. Although the Court of Appeals in Chance accepted the "strong showing" requirement as being consistent with the traditional rational relationship test,* the Court of Appeals for the First Circuit in Castro v. Beecher** took issue with this assumption and clearly stated that a new "substantial relationship" standard had to be applied. The next year, noting that since the original decision in Chance case law developments (primarily in the sex discrimination area) had been such that a "viable middle-ground test had in fact emerged," the Second Circuit joined the First Circuit in specifically enunciating a new standard for equal protection analysis in employment

*458 F. 2d 1158, 1177.

**459 F. 2d 725, 735-6 (1972).

impression that the Court was not eager to declare sex discrimination one of its "strict scrutiny" categories (virtually assuring plaintiffs' success), but, on the other hand, sensitivity to the rising tide of sex discrimination problems could not be set aside lightly under the traditional rational relationship rubric. Thus, the Court began to articulate a middle-ground approach in these cases. Under this approach, both the actual purpose of the legislation or state action at issue and the "fit" between that purpose and the means chosen to accomplish it are subjected to judicial examination. Only when a "fair and substantial relationship" rather than a mere "rational relationship" between a challenged state policy and a legitimate state purpose is shown will the action be upheld.*

The lower federal courts which initially considered employment discrimination cases brought under the equal protection clause felt a similar compunction to articulate a new middle-ground standard that would allow for substantial, though not "strict," scrutiny of challenged testing practices. These courts faced a dilemma. Title VII and its specific discriminatory impact standards did not, at the time, apply to actions of local government agencies, and these cases therefore were

*For the classic analytic discussion of the new middle-ground approach, see Gunther, "Foreword: In Search of Evolving Doctrine on the Changing Court: A Model for a Newer Equal Protection," 86 Harv. L. Rev. 1 (1972). The strongest explicit judicial pronouncements on this issue have been seen by Justice Marshall in support of the middle-ground approach (see, e.g., Rodriguez, supra at 411 U.S. at 98-110; Massachusetts Board of Retirement v. Murgia 427 U.S. 307, 318-321 (1976) and by Justice Rehnquist in opposition (see Trimble v. Gordon, 420 U.S. 762, 777-786 (1977).

discrimination cases.* Similar middle-ground approaches were subsequently widely adopted by several other courts which considered test validation under general constitutional equal protection jurisdiction.**

In sum, then, the development of the middle-ground standard in the above line of cases was, to a large extent, a mechanism utilized to incorporate the Griggs "demonstrable relationship" requirement, and the specific EEOC Guidelines standards, under the rubric of the middle-ground "fair and substantial relationship" approach in discriminatory impact cases brought under the equal protection clause rather than directly under

---

*Bridgeport Guardians v. Bridgeport Civil Service Commission 482 F. 2d 1333, 1336-7 (2d Cir. 1973). In this decision, the Second Circuit was affirming Judge Newman's decision discussed in detail at p. [29] ff, supra. Judge Newman had specifically articulated the need for enunciation of a "middle-ground" approach to equal protection analysis. Referring to the district court opinion in Chance, he pointed out through particular examples, the distinction between the traditional rational relationship test and a middle-ground test in the employment discrimination context:

> But the District Court [in Chance], as I read its decision, did not say that use of the exam was irrational in the sense that no person could rationally believe that the exam was of any value in selecting personnel.... [C]ontent validity and other standards from the field of psychological testing may often be found to be inadequately met, but rarely can one say the standard is so inadequately met that the use of the exam could not rationally be thought useful for personnel selection. 354 F. Supp. at 787.

**See, e.g., Douglas v. Hampton, 512 F. 2d 976, 989 (D.C. Cir. 1975), Officers for Justice v. Civil Service Commission, 371 F. Sup. 1328, 1335-6 (N.D. Calif. 1973). See also Armstead v. Starkville Municipal Separate School District, 461 F. 2d 276, 279-80 (5th Cir. 1972), Georgia Association v. Nix, 407 F. Supp. 1102, 1108-9 (N.D. Ga. 1976).

Title VII.* Because of the Supreme Court's firm rejection of the prevalent assumption that Title VII standards could be fully read into general constitutional requirements in Washington v. Davis, the question naturally arises as to whether the Court also specifically meant to mandate a complete return to the traditional two-tier equal protection approach in this area.

Although the Court did not specifically focus its decision on this issue, its approach to the validity of Test 21 seems to indicate that more than a traditional "rational relationship" to any legitimate state purpose had to be shown. The court's reasoning can be summarized as follows:

> Test 21, which is administered generally to prospective Government employees, concededly seeks to ascertain whether those who take it have acquired a particular level of verbal skill; and it is untenable that the Constitution prevents the Government from seeking to modestly upgrade the communicative ability of its employees rather than to be satisfied with some lower level of competence particularly where the job requires special ability to communicate orally and in writing.**

The Court's brief analysis of the validity of Test 21 did not merely state, as a court would under traditional rational relationship standards, that testing was a reasonable practice for Civil Service hiring, and absent a showing of bad faith the court would defer to the expertise of the Civil Service administrators. Instead, the court could be said to have engaged in an independent (though brief) analysis of the test, an analysis

---

*See, e.g., Officers for Justice, supra at 1336, Castro v. Beecher, supra at 733.

**426 U.S. at 245-246.

which satisfied it that the purposes behind the testing requirement were substantial and the "fit" between the means chosen and the valid purpose was adequate.*

In short, although the extent of judicial scrutiny of the testing practices at issue here was clearly less than the probing analysis that would be required under the EEOC standards, the defendants were called upon to justify their practices in substantive terms. Washington v. Davis, thus may be interpreted as implicitly endorsing a middle-ground approach in this area, albeit a more "flexible" middle-ground approach than had been used by the lower federal courts previously.**

The only reported post-Davis decision to have directly considered this issue so interpreted the implication of Davis

---

*The Court also indicated that the district court had considered in detail the relevance of Test 21 to the training regimen (Id. at 252-3) and Justice Stevens additionally focused on both the manifest relevance of reading ability to the police function and the familiarity of judges with the needs of the law enforcement profession. (Id. at 254-255). Compare, however, the conclusory interpretation of Davis as applying a traditional rational relationship test set forth in "Adoption of Questions and Answers to Clarify and Provide a Common Interpretation of the Unifrom Guidelines on Employee Selection Procedures" 43 Fed. Regs. 11995, 12002. (Mar. 2, 1979).

**The Court's main concern in its lengthy discussion of the intent/impact issue was apparently to ensure that the state would not be required to prove a "compeling justification" for a practice having only discriminatory impact (Id. at 248). In other words, its focus was to clarify the inapplicability, absent a showing of discriminatory intent, of the rigorous "compeling state interest test" rather than on discussing precisely which "lesser" standard should apply.

on this point.* In Richardson v. McFadden** the Court of Appeals for the Fourth Circuit, after holding that in the absence of discriminatory intent, Title VII standards would not apply, specifically stated:

> We agree with the Fifth Circuit in Tyler v. Vickery...that under the Equal Protection Amendment the issue is still whether the examination is job-related, albeit a less demanding inquiry. The hallmark of a rational classification is not merely that it differentiates, but that it does so on a basis having a fair and substantial relationship to the purpose of the classification. Id at 1099. And here the purpose of the classification is to distinguish between persons demonstrating minimal competence to practice law and those lacking such knowledge and skill.***

*In a post-Davis appeal of certain remedial issues in Chance v. Board of Examiners, 561 F. 2d 1079 (2d Cir. 1977), plaintiffs specifically argued that Davis did not overrule the substantive equal protection analysis adopted by the court in the earlier stages of the case. The appeals court, however, decided the immediate issues without reaching this question.

**540 F. 2d 744, 748 (1976).

***As the court noted in McFadden, the decision of the Fifth Circuit in Tyler, the single court of appeals decision to have adopted the discriminatory intent standard prior to Davis, was cited by the Supreme Court with approval in that case. The applicability of the "fair and substantial relationship test" was also discussed by the three-judge court in South Carolina, supra, which although perhaps otherwise inclined to apply the traditional rational relationship approach, apparently felt constrained by the Fourth Circuit's opinion in McFadden to analyze the facts under a "fair and substantial relationship" approach.

Note that in applying the middle-ground test, the court in McFadden held that the challenged bar examination was fairly designed to test minimal competence in the field of law. In considering the precedential impact for other certification and licensing situations of this and other cases upholding the bar examination, the fact that judges, as lawyers, tend to be personally familiar with these tests (and, in addition, would have personally performed well on them) might be of some relevance.

Thus, although the impact of Davis on this issue is still not fully clear, the middle-ground "fair and substantial" relationship approach to equal protection analysis, which has been extensively applied in cases involving employment selection testing, would appear to be a viable approach for judicial consideration of other testing practices.*

*Virtually all of the cases invoking the middle ground equal protection test to date, both in the employment discrimination and in the educational testing contexts, have involved situations of discriminatory impact against racial minority groups. (Cf. Debra P. v. Turlington, p. [87] ff. supra, where claims of a class encompassing all students, white and black, were considered simultaneously with claims of specific subclasses of black students.) The reality of such impact, in fact, was what induced many courts to search for a means for giving substantial consideration to minority plaintiffs' claims, even though in the absence of specific discriminatory intent, "strict scrutiny" could not be applied. The logic of the "fair and substantial" relationship test is such, however, that (as in Debra P.) non-minority plaintiffs in future cases where a clear pattern of unvalidated testing practices is established may also be held by the courts to be entitled to relief under the equal protection clause. Note in this regard that in Davis, the Supreme Court indicated that absent a showing of discriminatory intent, a plaintiff "whether white or black" (Id. at 245) would be testing practices. In other words, the emphasis on discriminatory impact entitles minority plaintiffs to no special consideration. Hence, it may be that after Davis, either all plaintiffs in testing situations will be afforded middle-ground scrutiny, or none will.

Thus far, it might be said that invocation of the middle-ground test has generally been triggered when plaintiffs constituted a "semi-suspect" class (i.e., sex discrimination cases) or a suspect class not entitled to full strict scrutiny treatment (i.e., blacks establishing discriminatory impact, but not discriminatory intent in employment practices). The reading of Davis-McFadden set forth above might imply an additional broader "triggering" concept calling for invocation of middle-ground analysis in certain substantive areas which through historical development or as a matter of judicial policy justify a moderate degree of judicial scrutiny, whether or not "suspect" or "semi-suspect" groups constitute the plaintiff class. Employment selection and certain educational testing practices may fall into this category.

FOOTNOTES

1. Fincher, "Personal Testing and Public Policy," 28 Am. Psych. 489, 494 (1973). Similarly, a body composed of representatives from industry associations (such as the International Personnel Management Association), called the Ad Hoc Group on the Proposed Uniform Guidelines on Employee Selection Procedures, asserted to the EEOC in its "Comments" dated February 17, 1978 (at p. ii):

> The professional community over the last several years has made known its views that these guidelines presented a set of requirements with which no employer, even the largest, and no professional, even the most expert, could comply.

2. See, e.g., Lerner, "Washington v. Davis: Quantity, Quality and Equality in Employment Testing," 1976 Sup. Ct. Rev. 263, 304; Johnson, "Albemarle Paper Company v. Moody: "The Aftermath of Griggs and the Death of Employee Testing," 27 Hast. L. J. 1239 (1976); Note, "Developments in the Law: Employment Discrimination and Title VII of the Civil Rights Act of 1964," 84 Harv. L. Rev. 1109, 1127 (1971); Wilson, "A Second Look at Griggs v. Duke Power Co.: Ruminations on Job Testing, Discrimination and the Role of the Federal Courts," 58 Va. L. Rev. 844 (1972); N. Glazer, Affirmative Discrimination 56 (1975).

3. See Chapter 1 supra.

4. For example, the "Standards for Educational and Psychological Tests and Manuals" published by the American Psychological Association have repeatedly been considered and cited by the courts in these cases. The original Guidelines on Employee Selection Procedures issued by the Equal Employment Opportunity Commission, 29 C.F.R. Part 1607 (1970) [hereinafter referred to as the "1970 EEOC Guidelines"], as well as later guidelines adopted by the EEOC and other federal agencies on which the courts have extensively relied, also incorporate these standards. In fact, the original version of the EEOC Guidelines was largely generated by a group of testing experts brought together for that purpose by the EEOC's office of research. Blumrosen, "Strangers in Paradise: Griggs v. Duke Power and the Concept of Employment Discrimination," 71 Mich. L. Rev. 59, 97 (1972). And, the recently-promulgated 1978 Uniform Guidelines jointly adopted by the EEOC and other federal executive agencies were enacted only after receipt, and consideration, of voluminous written and oral testimony by individual psychologists and official representatives of the APA, the International Personnel Management Association and more from 200 other groups.

Uniform Guidelines on Employee Selection Procedures (1978), 43 Fed. Reg. 38, 290, 38 295 (Aug. 25, 1978) [hereinafter referred to as "the 1978 Uniform Guidelines"]. See also Proposed Rulemaking on Uniform Guidelines on Employee Selection Procedures 42 Fed. Reg. 65541 (Dec. 30, 1977). Gorham, "Whom Does Government Listen To?" 29 Pers. Psych. 530 (Winter, 1976). According to Prof. Alfred Blumrosen, a prime consultant to the EEOC during these deliberations, the main professional concerns centered on definitions of content and construct validity. Although the state of professional opinion was far from unanimous on the points raised, substantial modifications were made in the final draft of the 1978 guidelines in response to these concerns. (Interview, April 2, 1979)

5. For a discussion of criterion validation and alternative testing requirements, see p. [17], infra. However, although it has repeatedly been pointed out that the validity of the concept of differential validation for discrete minority groups has not been fully substantiated in the psychological literature (see, e.g., Johnson, supra note 2, at 1259. Note, "Proof of Job-Relatedness," 52 Notre Dame Law R. 95, 104 (1976). Note, "Developments in the Law: Employment Discrimination in Title VII of the Civil Rights Act of 1964," 85 Harv. L. Rev. 1109, 1128 (1971); Lerner supra note 2 at 293), the courts have nevertheless tended to insist on adherence to differential validation requirements in cases where a substantial adverse impact on minority job applicants has been established. See, e.g., United States v. Georgia Power Company, 474 F. 2d 906, 914 (5th Cir. 1973), Rogers v. International Paper Co., 510 F. 2d 1340, 1350 (8th Cir. 1975). See also Albermarle Paper Co. v. Moody 422 U.S. 405, 425 (1975), Cooper and Sobol, "Seniority and Testing under Fair Employment Laws: A General Approach to Objective Criteria of Hiring and Promotion," 82 Harv. L. Rev. 1598, 1645-6 (1969).

6. For a detailed discussion of the legitimacy and effectiveness of judicial involvement in public policy making, see M. Rebell and A. Block, Educational Policy Making and the Courts (forthcoming).

7. For a complete understanding of the state of the law on competency assessment issues, it would, of course, be necessary to also thoroughly survey the numerous rulings of the Equal Employment Opportunity Commission in its probable cause and conciliation decisions. Such an examination is beyond the scope of the present effort. EEOC decisions and conciliation agreements obviously have a substantial impact on practices in the field, but if a respondent is unwilling to accept the Commission's rulings, it can only enforce its orders by bringing the case to the courts. See 42 U.S.C.A. SS2000-e.5 et seq. For a review of EEOC decisions on some of the issues raised in this paper,

see Connolly and Connolly, "Equal Employment Opportunities: Case Law Overview," 29 Mercer L. Rev. 677 (1978). Note also that because the focus of the present chapter is on the judicial treatment of test validation issues, the review of the cases contained herein will not consider the additional, complex issues concerning the requisite degree of "adverse impact" necessary in most cases to trigger the detailed court scrutiny of testing practices under Title VII of the 1964 Civil Rights Act. In other words, in almost all of the cases discussed in this chapter, the courts will have found a pattern of adverse impact as a threshold matter before commencing their discussion of the validation issues.

8. Cooper and Sobol, supra note 5, at 1637; "Special Issue: testing and Public Policy," 20 Amer. Psych. 857 (Nov. 1965). As one commentator put it, "there are more ability tests being given annually in the United States than there are people." Note, "Developments in the Law: Employment Discrimination and Title VII of the Civil Rights Act of 1964," 84 Harv. L. Rev. 1109, 1120 (1971).

9. Myart v. Motorola Company, reprinted in 110 Cong. Rec. at 5662.

10. 42 U.S.C. § 2000-s(h).

11. As indicated in note 7, specific enforcement of the provisions of the act could only be undertaken through the courts.

12. Blumrosen, supra, noe 4, at 97. See also Note, "Employment Testing and FEA Guidelines on Employment Selection Procedures: One Step Forward and Two Steps Backward for Equal Employment Opportunity," 26 Cath. L. Rev. 852 (1977).

13. Albermarle v. Moody, 422 U.S. 405, 452 (1975), (Burger, C.J., dissenting). See Note, "Weight of EEOC Guidelines in Evaluation of Employment Selection Procedures," 50 Tul. L. Rev. 397 (1976).

14. 296 F. Supp. 40 (N.D. Ala. 1968).

15. 2 EPD S10, 293 (N.D. Ga. 1970).

16. See also Broussard v. Schlumberger Well Services, 315 F. Supp. 506, 512 (S.D. Texas 1970).

17. 296 F. Supp. at 76.

18. 292 F. Supp. 413 (S.D. Ohio 1968).

19. 306 F. Supp. 1355 (D. Mass. 1969).

20. 319 F. Supp. 314 (E.D. La. 1970).

21. See also Penn v. Stumpf, 308 F. Supp. 1238 (N.D. Calif. 1970)(motion to dismiss denied in case involving mental ability and general tests for a policeman's job).

22. Under some circumstances transfer of incumbent employees to a more desirable job level would be permitted if one, rather than both, of the prerequisites were met.

23. Griggs v. Duke Power Company, 292 F. Supp. 243 (M.D.N.C. 1968).

24. Griggs v. Duke Power Company, 420 F. 2d 1225, 1234 (4th Cir. 1970).

25. Griggs v. Duke Power Company, 401 U.S. 424 (1971).

26. Id. at 436.

27. Id. at 431.

28. Id. at 433-434.

29. United States v. Georgia Power Company 474 F. 2d 906, 913 (5th Cir. 1973). Some courts went further, indicating that the guidelines "control." See, e.g., Rogers v. International Paper Company, 510 F. 2d 1340, 1345 (8th Cir. 1975). Many courts made strict compliance with the Guidelines a requirement in their remedial decrees. See, e.g., EEOC v. Detroit Edison Co., 515 F. 2d 301, 317 (6th Cir. 1975).

30. See, e.g., Vulcan Society of the N.Y.C. Fire Department v. Civil Service Commission, 360 F. Supp. 1265 (S.D.N.Y. 1973), aff'd, 490 F. 2d 387 (2d Cir. 1973); Douglas v. Hampton, 512 F. 2d 976 (D.C. Cir. 1975), Officers for Justice v. Civil Service Commission, 371 F. Supp. 1328 (N.D. Calif. 1973), U.S. v. Local 638, 360 f. Supp. 979 (S.D.N.Y. 1973), modified in part, 501 F. 2d 622 (2d Cir. 1974).

31. 422 U.S. 405 (1975).

32. Id. at 408.

33. Id. at 431.

34. Id. at 431.

35. Id. at 449 (Blackmun, J. concurring).

36. Id. at 451.

37. Cf. Note, "Application of EEOC Guidelines to Employment Test Validation," 41 G. W. U. L. Rev. 505 (1973).

38. This total consists of all cases annotated under the "Measuring Employment Qualifications" (SS460-484) section of the CCH Employment Practices Guide for the period in question, supplemented by additional test validation cases listed under head notes 49, 82 and 129 in the annotations to 42 U.S.C.A. S2000 e-2. (The lower court decisions in Albermarle and in Washington v. Davis were excluded from the sample.) This grouping, although incorporating all cases considered to be of significance by two independent annotators, of course, contains only reported decisions and by its very nature would tend to emphasize cases in which employment validation and testing issues were considered in substantive detail by the courts. Cases which were settled prior to issuance of a decision, or in which issues were decided in a summary manner, would, therefore, tend to be excluded from the sample.

39. 331 F. Supp. 1134 (S.D. Ala. 1971) aff'd 466 F. 2d. 122 (5th Cir. 1972). But see also, reaffirmance of original decision in Colbert, supra, note 15, at 4 E.P.D., S7780 (N.D. Ga. 1971), after remand following Griggs.

Allen involved achievement tests for transfer and promotion within a local police force. The test at issue was the "only known standardized police test," although it had not been validated for local use. Acknowledging that "comprehensive statistical studies" had not been undertaken, the court nevertheless refused to strike down the test because it seemed facially job-related and because invalidation of the test might open the door to a "spoils system" or at the least would require the city to spend an estimated $30,000 to conduct a test study.

Note, however, that in 1978 the same judge, ruling on a follow-up motion in the same case, held that in the light of the post-1972 applicability of Title VII to local governmental agencies, the emphasis on the EEOC Guidelines in Albermarle and the further precedents in the 5th Circuit, he would apply specific requirements of the new 1978 Uniform Guidelines, and award substantial relief to the plaintiffs. The fact that not one black had been promoted to sergeant during the 7 year interval between the two decisions also obviously influenced this result. Allen v. City of Mobile, 464 F. Supp. 433 (S.D. Ala. 1978).

40. Spurlock v. United Airlines, Inc. 330 F. Supp. 228 (D. Colo. 1971), aff'd, 475 F. 246 (10th Cir. 1972) (flight time requirements held to be substantially correlated with job performance by airline flight officials); Sims v. Sheet Metal Workers 353 F. Supp. 22 (N.D. Ohio 1972), aff'd in part, remanded in part, 489 F. 2d 1023 (6th Cir. 1973) (reading,

mathematics, and mechanical comprehension tests found to be validated, including differential validation); Kinsey v. Legg, Mason & Co., 8 E.P.D. 9767 (D.C., D.C. 1974) (general aptitude and sales batteries for security salesman found to have .49 correlation to valid criterion measures); Afro-American Patrolman's League v. Duck, 366 F. Supp. 1095 (N.D. Ohio 1975) (police exam held to be content and construct valid; performance evaluation alternative held "impractical"); Smith v. St. Louis Railway, 397 F. Supp. 580 (N.D. Ala. 1975) (general aptitude tests held job-related in accordance with Guidelines).

41. See, e.g., Hester v. Southern Railway Co. 497 F. 2d 1374 (5th Cir. 1974) (discriminatory impact not established), Frontera v. Sindell, 522 F. 2d 1215 (6th Cir. 1975) (no requirements to give exam in Spanish), Tyler v. Vickery, 517 F. 2d 1089 (5th Cir. 1975) (discriminatory intent required to invoke Title VII standards in challenge to bar examination), Mele v. U.S. Dep't. of Justice, 395 F. Supp. 592 (D. N.J. 1975) (Title VII protection held not to apply for white male).

42. See, e.g., Chance v. Board of Examiners, 330 F. Supp. 203 (S.D. N.Y. 1971) aff'd 458 F. 2d 1167 (2d Cir. 1972); Smith v. City of E. Cleveland, 363 F. Supp. 1131 (E.D. Ohio 1973), rev'd in part, 520 F. 2d 492 (6th Cir. 1975), Officers for Justice v. Civil Service Commission, 371 F. Supp. 1328 (N.D. Calif. 1973); Watkins v. Scott Paper Company, 530 F. 2d 1159 (5th Cir. 1976).

43. 490 F. 2d 387, 394 (2d Cir. 1973).

44. See, e.g., Vulcan, note 30 supra at 395, Douglas v. Hampton, 512 F. 2d 976, 984 n. 62 (D.C. Cir. 1975).

45. These impressions are consistent with the specific findings concerning the judicial capacity to engage in sophisticated social fact-finding as reported in Rebell & Block, supra, note 6.

46. See, e.g., Chance, note 42 supra; Vulcan, note 30 supra; but cf. Hampton, note 45 supra. In this regard, the courts anticipated the changes eventually incorporated into the 1978 Uniform Guidelines which now acccept content, construct or criterion-related validation as "equivalent approaches."

47. 29 C.F.R. +1607.3. See Harper v. Mayor and City Council of Baltimore, 359 F. Supp. 1187, 1204-5 (D. Md. 1973) and cases cited therein. On this point the language in S3.B of the 1978 Uniform Guidelines substitutes exhortatory language for the original mandatory requirement in the EEOC Guidelines. See also Allen v. City of Mobile, 464 F. Supp.. 433 (S.D. Ala. 1978).

48. See, e.g., United States v. Jacksonville Terminal Company, 451 F. 2d 418 (5th Cir. 1971); Crockett v. Green, 388 F. Supp. 912 (E.D. Wisc. 1975), aff'd, 534 F. 2d 715 (7th Cir. 1976). The third of the highly controversial EEOC provisions, differential validation, was broadly invoked by the courts, as indicated at note 5 above, even when professional reservations on the practice were known apparently because such standards went to the heart of equal employment opportunity law enforcement. See Georgia Power, supra note 29, at 914.

49. 426 U.S. 229 (1976).

50. As the Court itself noted (426 U.S. at 245 n. 12) virtually every federal District Court and Court of Appeals which considered the issue, had assumed that the Title VII impact standard should be incorporated into constitutional employment discrimination cases, with the single exception of the Fifth Circuit's decision in Tyler v. Vickery, 517 F. 2d 1089 (1975). Not mentioned in that footnote is the fact that the Court itself had directly encouraged these understandings, both by refusing to accept certiorari and review the lower court decisions in those cases (See, e.g., Bridgeport Guardians v. Bridgeport Civil Servants Association, 354 F. Supp. 778 (D. Conn. 1973), aff'd in part, rev'd in part, 482 F. 2d 1333 (2d Cir., 1973), cert. den. 421 U.S 991 (1975)) and by its specific notation in McDonnell Douglas Corporation v. Green, 411 U.S. 792, 802, n. 14 (1973) which cited the decisions in Chance and Castro as being proper demonstrations of the application of the Griggs demonstrable relationship requirements. It would appear, then, that the Court itself had accepted the logic of the lower courts' incorporation of Title VII impact standards in constitutional employment discrimination cases, unitl it had occasion to fully focus on the implications of these development for broader fields of constitutional law. Since the 1972 amendments had extended Title VII's coverage to virtually the entire employment discrimination area, pronouncement of a new constitutional intent standard in Davis could correct some developing precedents in these cases, without substantially retarding enforcement of the actual job-relatedness requirements. Since Davis, several district courts have held that Title VII cases involving public sector defendants also would require an intent standard, but these decisions have been largely reversed by recent circuit court opinions which have held that the statutory impact standard applied to public as well as private employers. Scott v. City of Aniston, 430 F. Supp. 508 (N.D. Ala. 1977) rev'd. 597 F. 2d 897 (5th Cir. 1979); Blake v. City of Los Angeles, 435 F. Supp. 55 (C.D. Calif. 1977), rev'd 595 F. 2d 1367 (9th Cir. 1979). See also U.S. v. City of Chicago, 573 F. 2d 416 (7th Cir. 1978). Note, "Title VII and Public Employers: Did Congress Exceed Its Powers," 78 Col. L. Rev. 372 (1978). The Supreme COurt has avoided ruling directly on this issue in

several cases whre it was raised. See, e.g., Hazelwood School District v. United States, 433 U.S. 299, 306 n. 12 (1977), New York City Transit Authority v. Beazer, 99 S. Ct. 1355, 1365 n. 25 (1979). See also Fitzpatrick v. Bitzer 427 U.S. 445, 456, n. 11 (1976), County of Los Angeles v. Davis 99 S. Ct. 1379 (1979) (case declared moot, analogous S1981 intent issue not reached).

51. [M]any observers believe the court in Washington v. Davis backed away from any reference in Albermarle that strict adherence to the guidelines is necessary." Connolly and Connolly, "Equal Employment Opportunities, Case Law Overview," 29 Mercer L. Reve. 637, 700 (1978). See also Image of Greate San Antonio v. Brown 570 F. 2d 517, 520 (5th Cir. 1978).

52. Davis v. Washington , 348 F. Supp. 15 (D. D.C. 1972).

53. Washington v. Davis, 512 F. 2d 956 (D.C. Cir. 1975).

54. 426 U.S. at 250-251 (emphasis added).

55. 433 U.S. 321 (1977).

56. See, e.g., McDonald v. Santa Fe Rail Transportation Company, 427 U.S. 273, 279 (1976); Nashville Gas Company v. Satty, 434 U.S. 136, 142-43 (1977); TWA v. Hardison, 432 U.S. 63 (1977). See also N.Y.C. Transit Authority v. Beazer 99 S. Ct. 1355, 1366 n. 31 (1979). In U.S. Board of Commissioners, 435 U.S. 110, 134 (1978), after citing Albemarle, the Court held that:

> When a Congress that reenacts a statute voices its approval of an administrative or other interpretation thereof, Congress is treated as having adopted that interpretation and this court is bound thereby.
>
> The Legislative history of the 1972 amendments to Title VII clearly indicates that both the Griggs decision and the specific EEOC guidelines were understood and approved. See Hourse Report 92-238. See also Lorillard v. Pons, 434 U.S. 575, 580 (1978).
>
> Note, however, that where the court believes that a position of an administrative agency, such as the EEOC, is inconsistent with legislative intent, its guidelines will not be followed. See, e.g., General Electric Company v. Gilbert, 429 U.S. 125 (1976), but cf. Nashville Gas Co., supra.

57. This post-Davis grouping consists of all cases decided after the fall of 1976, and annotated under headnotes 49, 82, 82-a and 129 in the annotations of 42 U.S.C. S2000e.2 through

December, 1978, supplemented by shepardizing Davis under headnotes 9-16, through January 1979, for cases involving test validation issues.

58. 549 F. 2d 415, 430 (7th Cir. 1977).

59. Similarly, the Eighth Circuit in Donnell v. General Motors Corporation, 576 F. 2d 1292 (8th Cir. 1978), stated that the burden to show job relatedness continued to be "heavy." Accord, Black v. City of Los Angeles, 595 F. 2d 1367 (9th Cir. 1979); Davis v. County of Los Angeles, 566 F. 2d (9th Cir. 1977), dismissed as moot, U.S. (1979). Judge Newcomer, in Dickerson v. U.S. Steel, 17 E.P.D. S8525 (E.D.P. 1978), in a voluminous opinion, undertook an Albermarle-type approach in measuring defendant's test validation attempts against each of the specific relevant guideline standards; he concluded that they were wanting in most respects. Accord, Vanguard Justice Society Inc. v. Hughes, 471 F. Supp. 670 (D. Md. 1979).

60. 464 F. Supp. 433 (S.D. Ala. 1978).

61. However, in situations involving very simple facts, some courts have basically ignored the EEOC guidelines and accepted the facial validity of the challenged practice. See e.g., Jenkins v. Caddo-Bossier Association for Retarded Children, 570 F. 2d 1227 (5th Cir. 1978) (unvalidated promotion standards for employees of small, sheltered workshop for mentally handicapped involving 26 employees upheld on the basis that supervisors obviously knew all individuals and administration of objective criteria would be "impractical and foolish." Note also the recent shift inits position on this issue by the Fifth Circuit in James v. Stockham Valves and Fittings Co., 5659 F. 2d 310 (5th Cir. 1977), where the Fifth Circuit expressed its continued reliance on the Guidelines. A year later, however, in Image of Greater San Antonio v. Brown, 570 F. 2d 517 (5th Cir. 1978), the same court strongly questioned the validity of the specific "Griggs standards" in light of Davis, and indicated that a more general "good reason" standard might now be applicable.

62. 354 F. Supp. 778 (D. Conn. 1973).

63. Id. at 790.

64. Id. at 788.

65. 354 F. Supp. at 792,

66. 482 F. 2d 1333, 1337 (2d Cir. 1973).

67. 431 F. Supp. 931 (D. Conn. 1977). ("Bridgeport II")

68. 431 F. Supp. at 936.

69. 431 F. Supp. at 937.

70. Dr. Barrett is known to have testified in at least 54 testing cases as of 1977 and is reported to have concluded that the tests were not properly validated in 53. See Guardians Association of New York Police Department v. Civil Service Commission, 431 F. Supp. 536, 544 (S.D.N.Y.), vacated and remanded, 562 F. 2d 38 (2d Cir. 1977), on remand, 466 F. Supp. 1273 (S.D.N.Y. 1979)

71. Firefighters Institute for Racial Equality v. City of St. Louis, 410 F. Supp. 948 (E.D. Mo. 1976), aff'd in part, rev'd in part, 549 F. 2d 506 (8th Cir. 1977), cert denied, 434 U.S. 819 (1977) provides a good illustration of a post-Davis case in which plaintiffs prevailed. The fire captain's promotional exam at issue in that case had been constructed on the basis of a professionally-developed job analysis and test construction procedure somewhat similar to that approved by the court in Bridgeport II.

The appeals court held, however, tha although the test preparation here may have "appeared impressive in relation to those challenged in other cases" it still did not meet professional and statutory requirements. The fatal flaw in the validation process in this situation was that:

> The test Dr. O'Leary devised did not reflect his findings in the job analysis. The captain's exam admittedly failed to test the one major job attribute that separates a firefighter from a fire captain, that of supervisory ability. From the interviews conducted for the job analysis, the City's expert determined that almost 43% of a fire captain's time was spent in supervision, a higher percentage of time than on any other single element. 549 F. 2d at 511

In other words, failure to test for an essential attribute of the job was a violation of +1607.5(a) of the EEOC Guidelines, as well as of the APA standards, both of which were specifically cited by the court.

Noting that there may be no good pen and paper test for evaluating supervisory skills, the court in Firefighters Institute went on to discuss the assessment center technique as one "excellent method of supervisory evaluation" which the parties might consider at a remedial stage. (Apparently, because of the high costs involved--estimated at as much as $500 per person--the court was not inclined to specifically mandate utilization of this particular approach). Its specific order was that the defendants must review their practices, and the "final test must be validated in accordance with the published EEOC guidelines."

72. See, e.g., Larry P. v. Riles, 343 F. Supp. 1306 (N.D. Calif. 1972), aff'd 502 F. 2d 963 (9th Cir. 1974); F. Supp. (N.D. Calif. Oct. 16, 1979)1; Copeland v. School Board, Portsmouth, Virginia, 464 F. 2d 932, 934 (4th Cir. 1972); Sorgen, "Testing and Tracking in Public Schools," 24 Hast. L.J. 1129, 1172 (1973). Cf. United States v. Norcome, 375 F. Supp. 270, 286 (D. D.C. 1974), aff'd, 497 F. 2d 686 (D.C. Cir. 1974).

73. See, e.g., Debra P. v. Turlington, F. Supp (M.D. Fla. July 12, 1979); see also articles cited in note 163 infra.

74. Morgan v. Kerrigan, 530 F. 2d 401, 414 (1st Cir. 1976); Board of Regents v. Bakke, 438 U.S. 265 (1978) (Brennan, J. concurring in part, dissenting in part).

75. See, e.g., 20 U.S.C. +1401 et seq. The provisions of 20 U.S.C. +1412(5) specifically require that testing and evaluation methods and procedures for the classification of handicapped children be selected and administered "so as not to be racially or culturally discriminatory." At least one commentator has interpreted these provisions as calling for the application of Griggs standards in this context. Note, "Enforcing the Right to an Appropriate Education: The Education for all Handicapped Children Act of 1975," 92 Harve. L. Rev. 1103, 1117 (1979). See also discussion of Larry P., pp. [72-746], infra.

76. See, e.g., 20 U.S.C. S3169c(1) (C); 45 C.F.R. S185.43.

77. See generally Kirp, "Law, Politics and Equal Educational Opportunity: The Limit of Judicial Involvement," 47 Harv. Ed. Rev. (1977).

78. 401. U.S. at 431.

79. 401 U.S. at 433.

80. See White and Francis, "Title VII and the Masters of Reality: Eliminating Credentials in the American Labor Market," 64 Geo. L. J. 1213, 1217-19 (1976); Note, "Diplomas, Degrees and Discrimination," 26 Hast. L. J. 1377 (1975)

81. See 29 C.F.R. +1607.2. Compare 1978 Uniform Guidelines S16.Q.

82. 401 U.S. at 433, no. 8.

83. Compare in this regard the position articulated by the Fifth Circuit:

> Graduation from high school demonstrates an ability to pass the various tests administered by the particular school.

> Certainly, where so many kinds of employment tests have been shown to be unrelated to job performance, it is likely that high school tests, not even designed to predict performance, are subject to the same flaws. Watkins v. Scott Paper Company, 530 F. 2d 1159, 1182 (5th Cir. 1976). (footnotes omitted). See also Johnson v. Goodyear Tire and Rubber Company, 491 F. 2d 1364, 1371-2 (5th Cir. 1974).

with the view of the California district court in League of United Latin American Citizens v. City of Santa Anna, 410 F. Supp. 873, 901 (C.D. Cal. 1976):

> This court, therefore, is reluctant to accept the idea that education requirements must be empirically validated. To accept that concept would be to adopt the proposition that the empiricist's method of arriving at truth are the only acceptable ones...

84. The methodology for selecting these cases was similar to the methodology used in the sample of cases discussed at p. 23, supra. Reported federal court opinions were located primarily by shepardizing Griggs (through February 1979) for all subsequent cases involving the diploma isue; in addition, to assure completeness of the sample, additional cases were sought in the CCH Employment Practices Guide (headnote 437); and the United States Code Annotated (headnotes to 42 U.S.C.A. 2000e).

85. The 33 decisions were contained in 29 separate cases. That is, in four of the 29 cases the court issued distinct opinions affecting two different job categories and qualification requirements.

86. Dobbins v. Electrical Workers Local 212, 292 F. Supp. 413 (S.D. Ohio 1968); Broussard v. Schlumburger Well Services, 315 F. Supp. 506 (S.D. Tex. 1970). In Dobbins, the court found in favor of the plaintiff on most claims, but upheld the high school diploma requirement without discussion. In Broussard, the court followed the lower court decision in Griggs (subsequently reversed by the Supreme Court) in striking down a diploma requirement as applied to blacks hired for lower level jobs at a time when whites could attain the higher level positions at issue without a diploma, but permitted continued use of diplomas for all new employees.

87. See, e.g., Duhan v. Goodyear Tire and Rubber., 494 F. 2d 817 (5th Cir. 1974) (tire factory); Padilla v. Stringer, 395 F. Supp. 495 (D.N.M. 1974) (zookeeper).

88. See, e.g., E.E.O.C. v. Local 638, 532 F. 2d 821 (2d Cir. 1976); Donnell v. General Motors Corp., 576 2d 1292 (8th Cir. 1978).

89. See, e.g., Carter v. Gallagher, 452 F. 2d 315 (8th Cir. 1972); Dozier v. Chupka, 393 F. Supp. 836 (S.D. Ohio 1975).

90. 474 F. 2d 906 (5th Cir. 1973).

91. Id. Griggs, of course, provided a similar factual situation. See 401 U.S. 424, 431 (1973). See also Pettway v. American Cast Iron Pipe Co., 494 F. 2d 211 (5th Cir. 1974); Dozier v. Chupka, 395 F. Supp. 836 (S.D. Ohio 1975).

92. Consideration of less discriminatory alternatives was specifically required underS+1607.3 of the 1970 EEOC Guidelines, and S3.B of the 1978 Uniform Guidelines. The Supreme Court in Griggs, interpreting Title VII and without specifically referring to the E.E.O.C. Guidelines on this point, emphasized that in this context "the touchstone is business necesssity." 401 U.S. at 431.

93. cf. Dothard v. Rawlinson, 433 U.S. 321 (1977) (test for "strength" recommended instead of height and weight requirements). In taking this position, the courts generally did not confront the further issue of whether such ability tests were themselves job-related.

94. Castro v. Beecher, 459 F. 2d 725 (1st Cir. 1972); Arnold v. Ballard, 390 F. Supp. 723 (N.D. Ohio 1975); League of United Latin American Citizens v. Santa Anna, 410 F. Supp. 873 (C.D. Cal. 1976); Morrow v. Dillard, 12 E.P.D. S11, 199 (S.D. Miss. 1976) on remand from 491 F. 2d 105 (5th Cir. 1974) cert. den. 419 U.S. 895 (1974); U.S. v. City of Buffalo, 457 F. Supp. 512 (W.D. N.Y. 1978). The diploma requirement in Washington v. Davis apparently was not challenged by plaintiffs; in any event, it was not presented for review by the Supreme Court.

95. This approval, of course, implicitly rejected the position taken by the E.E.O.C. and the Fifth Circuit Court of Appeals that diplomas should be subject to the same empirical validation standards as other tests. In two cases, this rejection was explicit. See Buffalo, supra, 457 F. Supp. at 628-629; and League, supra, 410 F. Supp. at 901. Note also the common judicial assumption that the functions of a police officer are "not foreign to judicial experience." See discussion at p [23], supra.

96. 457 F. Supp. at 629.

97. In U.S. v. Lee Way Motor Freight, Inc., 7 E.P.D.

S9066 (W.D. Okla. 1973), the supervisory positions at stake traditionally had been filled by persons working their way up from truck driver and foreman jobs; indeed a ranking supervisor

at time of trial was a former driver with a tenth grade education. In Payne v. Travenol Laboratories, 416 F. Supp. 248 (N.D. Miss. 1976), aff'd 565 F. 2d 895 (5th Cir. 1978), the court particularly noted that while the defendant required that "traffic analysts", "scheduling analysts" and "systems analysts" have college degrees, it specified no discipline or major area of study.

98. 475 F. 2d 216 (10th Cir. 1972). The soundness of the court's application of this rule to the evidence actually presented in Spurlock is questioned in Note, Employment Discrimination -- Building Up The Headwinds, 52 N. CAR. L. REV. 181 (1973).

99. Similar factors were mentioned in Rice v. City of St. Louis, 464 F. Supp. 138 (D. Mo. 1978) (public health representative) and in Morrow v. Dillard, note 91, supra. (state narcotics agent).

100. 475 F. 2d at 219.

101. 455 F. Supp. 1102 (D. Del. 1978), aff'd in part, rev'd in part, 601 F. 2d 76 (3rd Cir. 1979), cert. denied, U.S. (1979), 21 E.P.D. SS30, 316. In its partial affirmance, the Third Circuit Court of Appeals ruled that the plaintiff was not an adequate representative of persons who had been denied teaching positions by application of the degree requiremnt since he himself held a doctorage degree at the time he was hired. Consequently, it reversed that part of the district court decision that had certified a class action, and had named Scott as class representative. Thus, the district court's holding on the degree requirement issue must be now considered as dictum."

102. Interestingly, the E.E.O.C. (participating as an amicus curiae) initially argued that empirical validation was not necessary in this context but later retreated to a neutral position. 455 F. Supp. at 1125 n. 64.

103. 455 F. Supp at 1126.

104. This latitude is also evident in the diverse holdings in an analogous body of cases, those involving challenges to other non-test employment criteria such as height, weight, experience or personal character requirements. The amount of psychometric analysis in these cases ranges from common sense findings that a good back is manifestly related to performing manual labor (Smith v. Olin Chemical Corp., 555 F. 2d 1283 (5th Cir. 1977) (en band), or that a driver's accident record is related to the job of truck driving (Adams v. Texas & Pacific Motor Transport CO., 408 Supp. 156 (E.D. La 1975)); to detailed consideration

of the minimum height of a commercial airline pilot (Boyd v. Ozard Air Lines, Inc. 568 F. 2d 50 (8th Cir. 1977); or of the relevance of recent legal writing experience for an attorney's position in a government agency (Coopersmith v. Roudebush, 517 F. 2d 818 (D.C. Cir. 1975)). While expert testimony is common in these cases, empirical validation studies infrequently appear, and the legal status of the E.E.O.C. Guideline requiring empirical validation for height and weight requirements (29 C.F.R. 1606.1(b)) has not been settled by the Supreme Court. However, the Supreme Court's decision in Dothard v. Rawlinson, discussed at p. [24-24] supra, probably will accelerate the trend toward first approaching these cases by applyign business necessity doctrine, calling for tests of strength rather than absolute height/weight requirements, thereby avoding empirical validation issues in many instances. (See also Crockett v. Green, 338 F. Supp. 912, 920-21 (E.D. Mo. 1975), aff'd, 534 2d 715 (7th Cir. 1976)).

105. Uniform Guidelines SS2. B, 15 Q; cf. 1970 E.E.O.C. Guidelines, S1607.2.

106. In fact, the Uniform Guidelines specifically state at +2.B that the guidelines apply to "licensing and certification boards, to the extent thatttlicensing and certification may be covered by Federal equal employment opportunity law."

107. See Tyler v. Vickery, 517 F. 2d 1089 (5th Cir. 1975); Woodward v. Virginia Board of Bar Examiners, 420 F. Supp. 211 (E.D. Va. 1976); Delgado v. McTighe, 442 F. Supp. 725 (E.D. Pa. 1977).

108. See, e.g., Vanguard Justice Society, Inc. v. Hughes, 471 F. Supp. 670 (D. Md. 1979), where the court held that

> The term "employer" as it is used in Title VII is sufficiently broad to encompass any party who significantly affects access of any individual to employment opportunities, regardless of whether that party may technically be described as an "employer" of an aggrieved individual as that term has generally been defined at common law. Despite its concededly limited role in the hiring process, the Baltimore Civil Service Commission exercised substantial authority and discretion in the area of testing of applicants for entry level positions with the department.

Id. at 696. See also Curran v. Portland Superintency School Comm., 435 F. Supp. 1067 (D. Me. 1977); Puntolillo v. New Hampshire Racing Commission, 375 F. Supp. 1089 (D.N.H. 1971); Sibley Memorial Hosp. v. Wilson 410 F. Supp. 513 (N.D. Cal. 1976).

109. See United States v. South Carolina, 445 F. Supp. 1094, 1110 (D.S. C . 1977), aff'd., 434 U.S. 1026 (1978); Chance v. Board of Examiners, 330 F. Supp. 203 (S.D.N.Y. 1971), aff'd., 458 F. 2d 1167 (2d Cir. 1972); United States v. North Carolina, 400 F. Supp. 343 (E.D. N. Car. 1975), vacated 425 F. Supp. 789 (E.D. N.C . 1977).

110. Baker v. Columbus Municipal Separate School District, 329 F. Supp. 706 (N.D. Miss. 1971), aff'd 462 F. 2d 1112 (5th Cir. 1972); Walston v. County School Board of Nasemond County, 492 F. 2d 919 (4th Cir. 1973), U.S. v. North Carolina 400 F. Supp. 343 (E.D.N. Car. 1975), vacated, 425 F. Supp. 789 (1977); Georgia Association v. Nix, 407 F. Supp. 1102 (N.D. Ga. 1976). See also Armstead v. Starkville Municipal Separate School District, 325 F. Supp. 560 (W.D. Miss. 1971), aff'd, 461 F. 2d 276 (5th Cir. 1962). Note that all of these cases took place in the deep South, and the cases reflected patterns of present purposeful discrimination, or lingering effects of such discrimination in the past.

111. 351 F. Supp. 196, 205.

112. 492 F. 2d at 926.

113. 445 F. Supp. 1094 (D.S.C . 1977), aff'd, 434 U.S. 1026 (1978).

114. 445 F. Supp. at 1112.

115. Both S1607.5(b)(3) of the E.E.O.C. Guidelines and S12(b)(3) of the FEA Guidelines required that if success in a training program were to be utilized as an employment selection measure, the job relevance of the training program needed to be established. Accord, 1978 Uniform Guidelines SS14.B(3), 14.C(6).

116. 445 F. Supp. at 1113.

117. 445 F. Supp at 1115-6.

118. 445 F. Supp. at 1108 n. 13.

119. Cf. Furnco Construction Corp. v. Waters 438 U.S. 567 (1978) where in declining to invalidate a highly questionable hiring procedure, the Court held that "there is nothing in the record to indicate that the proposed alternative method would be any less 'haphazard, arbitrary, and subjective.'" 438 U.S. at 578.

120. 434 U.S. at 1027.

121. 330 F. Supp. 203 (S.D.N.Y. 1971), aff'd, 458 F. 2d 1167 (2d Cir. 1972)

122. Firefighters Institute v. City of Saint Louis, 549 F. 2d 506, 513 (8th Cir. 1977) (assessment center considered a concededly good device" to be considered in remedial stage); E.E.O.C. v. duPont, 445 F. Supp. 223, 254 (D. Del. 1978 (in-basket assessment test conceded by plaintiffs to be job-related); cf. Friend v. Leidlinger, 446 F. Supp. 361, 377 (E.D. Va. 1977), aff'd, 588 F. 2d 61 (4th Cir. 1978) (claim that assessment center evaluators were unqualified not considered by the court because assessment process had no discriminatory impact).

123. No footnote

124. 422 U.S. at 433.

125. 457 F. 2d 348 (5th Cir. 1972).

126. Wade v. Mississippi, 372 F. Supp. 126, 142 (N.D. Miss. 1974), aff'd, 528 F. 2d 508 (5th Cir. 1976). Similarly, criteria such as "verbal expression, appearance, maturity, drive," were considered unacceptable, subjective criteria in Robinson v. Union Carbide Company, 538 F. 2d 562, 662 (5th Cir. 1976).

127. 438 F. Supp. 213, 220 (E.D. Pa. 1977)

128. See Comment, "Subjective Employment Criteria," 54 Detroit J. Urban L. 165 (1976). Note in this regard the emphasis upon "observable aspects of work behavior of the job" in +14.C. (4) of the 1978 Uniform Guidelines. See also SS14.A(2) and (3).

129. 435 U.S. 79 (1978).

130. 435 U.S. at 102.

131. ALthough the Horowitz decision was issued in 1978, the reluctance to intervenc in academic judgments articulated therein represented an authoritative restatement of the law on this issue which had previously been widely stated by other federal courts. See,, e.g., the following cases, all of which were cited by the Supreme COurt in Horowitz: Gaspar v. Bruton, 513 F. 2d 843 (10th Cir. 1976), Greenhill v. Bailey, 519 F. 2d (8th Cir. 1975); Mahavongsanan v. Hall, 529 F. 2d 448 (5th Cir. 1976). See also Hubbard v. John Tyler Community College, 455 F. Supp. 753 (E.D. Va. 1978), Sofair v. Upstate Medical Center, 44 N.Y. 2d 475 (1978).

132. Singleton v. Jackson Municipal Separate School System, 419 F. 2d 1211, 1219 (5th Cir. 1969), rev'd in part on other grounds, 396 U.S. 290 (1970). Courts enforcing this prohibition noted that ability grouping had not previously been used in black schools prior to desegregation and that no real attempt at compensatory education was being undertaken for the larger black population assigned to the lower tracts. MOses v. Washington Parish School Board, 330 F. Sup. 1340 (E.D. La. 1971), aff'd, 456 F. 2d 1285 (5th Cir. 1972), cert. denied, 409 U.S. 1013 (1972). See also Lemon v. Bossier Parish School Board, 444 F. 2d 1400, 1401 (5th Cir. 1971) (tracking prohibited regardless of validity of testing); McNeal v. Tate County School Board, 508 2d 1017 (5th Cir. 1975); United States v. Gadsen City School District, 508 F. 2d 1017 (5th Cir. 1978) (ability grouping would continue to be prohibited until a showing is made that assignments are not based on results of past segregation or will remedy such results).

133. 343 F. Supp. 1306 (N.D. Calif. 1972) (preliminary injunction decision, aff'd, 502 F. 2d 963 (9th Cir. 1974), F. Supp. (Oct. 16, 1979) (decision on the merits).

134. 29 C.F.R +1607.8. (Larry P. v. Riles Civ. No. 71-2270, slip op. at 69 (Oct. 16, 1979)). Application of this standard amounted to a rejection of defendants' argument that I.Q. tests were permissible because they were widely used throughout the education system.

135. 29 C.F.R. S1607.5(b)(5) Larry P. v. Riles, note 134 supra, slip op. at 72.

136. Larry P. v. Riles, note 134 supra, slip op. at 65-66.

137. Id. at 67.

138. C-70 RFP (N.D. Cal. 1973) Consent Agreements Reproduced in Harvard Center For Law And Education, Classification Materials 199 (1973) (June, 1976 Supp at 39-40).

139. 269 F. Supp. 401 (D. D.C. 1967), aff'd sub nom. SMuch v. Hansen, 408 F. 2d 175 (D.C. Cir. 1969).

140. 269 F. Supp at 484.

141. 45 C.F.R. +185.43(c) (1). See Board of Education, Cincinnati v. Department of H.E.W., 396 F. Supp. 203 (S.D. Ohio 1975), modified on other grounds, 532 F. 2d 1070 (6th Cir. 1976) (regulations held not arbitrary, and "merely declare existing law"0.

142. Note, "Equal Protection and Intelligence," 26 Stan L. Rev. 647, 648 n. 6 (1974).

143. Kirp, "Schools as Sorters: The Constitutional and Policy Implications of Student Classification," 121 U. Pa. L. Rev. 705, 757 (1973); Sorgen, "Testing and Tracking in Public Schools," 24 Hast. L. J. 1129, 1168-70 (1973); Shea, "An Educational Perspective of the Legality of Intelligence Testing and Ability Grouping," 6 J. L. & Ed. 137, 141 (1977); Swift, "Testing: Misclassification and Invasion of Privacy" in Nolpe, Current Legal Issues in Education (1976). See also United States v. Norcome, 375 F. Supp. 270, 286 (D. D.C. 1974), aff'd, F. 2d 686 (D.C. Cir. 1974).

144. 501 F. 2d 1264 (9th Cir. 1974).

145. In its 1979 decision, the court in Larry P. note 134 supra, specifically cited Berkelman and reiterated these distinctions. Slip op. at 96.

146. 501 F. 2d at 1267.

147. See discussion on this point in Kirp, supra note 143, at 755-758, indicating that school aptitude tests are probably accurate predictors of subsequent school success, and, in fact, probably have greater predictive validation than employment tests, and concluding that cultural bias could be eliminated only by revamping the entire school system. Cf. discussion of Larry P. pp. [72-74b] supra.

148. 401 F. Supp. 216, 242 (D. Mass. 1975), aff'd, 530 F. 2d 401, 424 (1st Cir. 1976).

149. Larry P. v. Riles, note 134 supra, slip op. at 95-6.

150. 501 F. 2d at 1267. See also De La Cruz v. Torney, 582 F. 2d 45, 59 n. 10 (9th Cir. 1978), reaffirming the holding of Berkelman on this issue.

151. 396 F. Supp at 238. The law review commentaries which have considered these cases generally agree that the holdings in the ability testing-tracking cases can only be consistently understood at involving explicit or implicit application of the middle ground test. See Note, "Equal Protection and Intelligence," 26 Stan. L. Rev. 647 (1974); Kirp, note 143 supra at 744; Sorgen, note 143 supra at 1154; Shea, note 143 supra 154. Compare Swift note 143 supra at 198-202.

152. See also Spangler v. Pasadena City Board of Education, 311 F. Supp. 501, 519 (C.D. Calif. 1970).

153. See, e.g., Murray v. West Baton Rouge Parrish School Board, 472 F. 2d 438, 444 (5th Cir. 1973) (lack of showing of validation of psychological and achievement tests in the case of a few individual children does not give rise to a constitutional deprivation); Parish v. National Collegiate Athletic Association, 361 F. Supp. 1220, 1226-7 (W.D. La. 1973), aff'd, 506 F. 2d 1028 (5th Cir. 1978) (judicial notice taken of "widespread use of SATs"). But cf. Copeland v. School Board, Portsmouth, Virginia, 464 F. 2d 932 (4th Cir. 1972) (case remanded to determine if ability tests at issue were "reliable", citing Griggs).

154. 60 Cal. App. 3d 814 (1976).

155. Id. at 824 (emphasis added).

156. 47 N.Y. 2d 440 (1979.)

157. No. 562 (Slip op. dated Dec. 17. 1979), reversing 64 app. Div. 2d 369 (2d Dept. 1978).

158. Id. at 443.

159. 47 N.Y. 2d at 444-445. Similarly, in Hoffman the court warned:

> In order to affirm a finding of liability inthese circumstances, this court would be required to allow the finder of fact to substitute its judgment for the professional judgment of the Board of Education as to the type of psychometric devices to be used and the frequency with which such tests are to be given. Slip op. at 5.

In Peter W. the court took note of widespread "public dissatisfaction" with the public schools but concluded:

> [The schools] are already beset by social and financial problems which have gone to major litigation, but for which no permanent solution has yet appeard. [citations omitted] The ultimate consequence, in terms of public time and money (if malpractice claims were allowed), would burden them--and society--beyond calculation.

60 Cal. App. 3d at 861

The distinction between injunctive relief and money damages also was a central concern for the federal trial court and the court of appeals in Frederick L. v. Thomas, 419 F. Supp. 960 (E.D. Pa. 1976), aff'd, 557 F. 2d 373 (3d Cir. 1977) (injunction and 578 F. 2d 513 (3d Cir. 1978) (damage claim). Although the federal court ordered a comprehensive injunctive remedy for policies and practices that violated the rights of learning

disabled children, it abstained from a decision on the money damage claim of the named plaintiff, Frederick L., so that this sensitive and difficult state tort law issue could first be presented to the state courts. In other cases, handicapped children who were totally excluded from public education have requested compensation for their education injuries, either in the form of educational services past the age of 21, or money payments. In at least one case, in-kind services were ordered (Mattie T. v. Holladay, No. DC-75-31-5 (D. Miss., Jan. 26, 1979)).

160. For example, the New York State Commissioner of Education promulgated a MCT program in September 1978, but several months later amended it to add a section on remedial instruction. N.Y. Comm. Regs. S103.2(c).

161. See, e.g., McClung, "Competency Testing: Potential for Discrimination," 11 Clearinghouse Rev. 439 (1977); McClung, "Are Competency Testing Programs Fair? Legal?," 11 Phi Delta Kappan 397 (Feb. 1978); McClung and Pullin, "Competency Testing and Handicapped Students, 12 Clearinghouse Rev. 922 (1978); Tractenberg, "Testing for Minimum Competency/A Legal Analysis," (unpublished, AERA Topical Conference Oct. 1978); Tractenberg, "Who is Accountable for Pupil Illiteracy?" (Unpublished, 1978 National Right to Read Conference, May 1978); Tractenberg, The Legal Implications of Statewide Public Performance Standards (unpublished, Sept. 1977; an article styled as a reply to McClung and Tractenberg is Getz and Glass, "Lawyers and Courts as Architects of Educational Policy: the Case of Minimal Competence Testing," High Sch. J. 181 (Jan. 1979); Clague, "Competency Testing and Potential Constitutional Challenges of Everystudent'," 28 Cath. U.L. Reve. 469 (1979): Lewis, "Certifying Functional Literacy: Competency Testing Programs: Legal & Educational Issues," 47 Ford. L. Reve. 651 (1979); See also several articles published in the May, 1978 issue of Phi Delta Kappan; Haney and Madaus, "Making Sense of the Competency Testing Movement," 48 Harv. Ed. Rev. 462 (1978).

162. Civil No. 78-892-Civl.-T.-C (M.D. Gla. July 12, 1979)

163. Id., slip op. at 13.

164. Id., slip op. at 30.

165. DeFunis v. Odegaard, 416 U.S. 312 (1974).

166. Regents of the University of California v. Bakke, 438 U.S. 265 (1978).

167. A. Sindler, Bakke, DeFunis and Minority Admissions 312 (1978). See also Tribe, "Perspectives on Bakke: Equal Protection, Procedural Fairness, or Structural Justice," Harv. L. Rev. 864, 864-5 (1978)

168. Sindler, note 169 supra, at chapter 7. The studies generally indicate that the undergraduate grade point average is a better predictor of success in medical school than the MCAT.

169. See, e.g., ALevy, Downstate Medical enter, 39 N.Y. 2d 326, 330 (1976) (1970 Report of Association bias on MCAT"); Bakke v. Regents of the University of California, 132 Cal. Reptr. 680, 714 (1976) (Tobriner J. dissenting ("'objective' academic credentials on which the school had largely relied in the past did not accurately predict such minority applicants' qualifications...").

170. See DeFunis v. Odegaard, 507 P. 2d 1169, 1172-73 (S. Ct. Wash. 1973); L. Tribe, American Constitutional Law 1048 n. 33 (1978). Reply Brief for Petitioner, in Bakke, at p. 3:

> It is accurate to say that a minority admissions program results in selecting for admission from among many fully-qualified candidates some fully-qualified minority applicants who would not have been chosen under earlier color-blind criteria selection.
>
> The vice of the general labels "better qualified" and "less qualified," is that they confuse qualification for medical education and the profession with selection for admission from among the fully qualified applicants, and they then go on to assume, contrary to fact, that there is some abstract and universal measure of who is "better qualified" for all purposes.

171. 416 U.S. at 328-329, 334-335.

172. 416 U.S. at 330.

173. 416 U.S. at 340.

174. Procedure at the College of Human Medicine, Michigan State University, described in Braverman, "Beyond Bakke," (Am. Jewish Committee 1978).

175. There were some indications in Bakke that the general antidiscrimination standards of Title VI of the 1964 Civil Rights Act could be read to require, upon a showing of disparate impact, evidence of test validation in accordance with Griggs standards. It also should be noted that the four

justices joining in Justice Brennan's decision extended the scope of the "middle ground" equal protection test by specifically stating that this approach should not be limited to sex discrimination cases but also should be utilized in situations of remedial racial classification. This "middle ground" approach was also explicitly adopted by the New York Court of Appeals in Alevy, supra. See also Ely, "Foreword: On Discovering Fundamental Values," 92 Harv. L. Rev. 5, 12 (1978).

As discussed in the preceding section in regard to minimum competency testing at the secondary level, the apparent inadequacies of the present standardized admissions tests, which affect not only on minority members, but also upon thousands of non-minority applicants in a highly competitive situation, may also give rise to a general, non-racially oriented claim under the middle ground test. As Justice Tobriner indicated in the California State Supreme Court' decision in Bakke, "As medical school admissions officials themselves acknowledge, these studies raise questions of the most serious order as to the propriety of the continuing use of traditional admission criteria." (132 Cal. Rep. at 714)

176. For a discussion of testing applications in these areas, see W. Schrader, ed., Measurement and Educational Policy (1979).

177. See, e.g., James v. Board of Education of the City of New York, 42 N.Y. 2d 357 (1977) (refusal to interfere with policy judgments of educational officials on test security and reliability issues).

178. See, e.g., Robinson v. Cahill, 62 N.J. 473, 515 303 A. 2d 273, 295 (1973); and N.J. Stat. Ann +18A: 71-5 enacted as a result thereof.

179. See, e.g., EEOC v. Detroit Edison, 515 F. 2d 301 (6th Cir. 1975), vacated and remanded on other grounds, 431 U.S. 951 (1977), Kirkland v. N.Y. State Department of Correctional Services, 520 F. 2d 420 (2d Cir. 1975), Fire Fighters Institute v. City of St. Louis, 549 F. 2d 506 (8th Cir. 1977), Allen v. City of Mobile, 464 F. Supp. 433 (S.D. Ala. 1978)

180. For a further detailed discussion of the implementation of the Chance decree see Rebell & Block, note 6 supra, ch. 6. For further illustrations of complex interplay between affirmative action quota considerations and scientific validation standards arising at the remedial stages of employment discrimination cases, see, e.g., Western Addition Community Service Organization v. Alioto, 369 F. Supp. 77 (N.D. Calif. 1973), aff'd, 514 F. 2d 542 (9th Cir. 1975) and Armstead v. Starkville Municipal Separate School District, 325 F. Supp. 560 (N.O. Miss., 1971), modified, 461 F. 2d 276 (5th Cir. 1972).